VOLUME 20 NUMBER 4 DECEMBER 1957

SOCIOMETRY

A JOURNAL OF RESEARCH IN SOCIAL PSYCHOLOGY

Articles by

Coleman, Katz and Menzel
Farber, Harlow and West
Stryker
Hall
Hammer and Ikle
Borgatta

Published by

THE AMERICAN SOCIOLOGICAL SOCIETY

SOCIOMETRY

EDITOR

LEONARD S. COTTRELL, JR.

ASSOCIATE EDITORS

R. FREED BALES
HERBERT BLUMER
JOHN A. CLAUSEN
LEON FESTINGER
NELSON N. FOOTE
HERBERT HYMAN
IRVING L. JANIS
FREDERICK MOSTELLER
THEODORE M. NEWCOMB
WILLIAM H. SEWELL

EDITORIAL CONSULTANTS

BERNARD M BASS
EDGAR F. BORGATTA
ORVILLE G. BRIM, JR.
LAUNOR F. CARTE
FRED E. FIEDLER
ROBERT N. FORD
NATHAN L. GAGE
HERBERT GOLDHAMER
LEO A. GOODMAN
ALEX INKELES
WILLIAM W. LAMBERT
PHILIP J. MCCARTHY
RICHARD T LAPIERE
STANLEY SCHACHTER
HERBERT A. SIMON
SHIRLEY A. STAR
FRED L. STRODTBECK
GUY E. SWANSON
E. PAUL TORRANCE
JOHN W. M. WHITING
ROBIN M. WILLIAMS, JR.
ROBERT F. WINCH

EXECUTIVE OFFICER

MATILDA WHITE RILEY

Published four times a year

March, June, September, December

by

THE AMERICAN SOCIOLOGICAL SOCIETY

at The Waverly Press, Mt. Royal and Guilford Aves., Baltimore 2, Md.

Sociometry was founded in 1937 by J. L. Moreno

Communications for the Editors and all manuscripts should be addressed to Leonard S. Cottrell, Jr., Russell Sage Foundation, 505 Park Avenue, New York 22, New York. Business communications should be addressed to Matilda White Riley, The American Sociological Society, New York University, Washington Square, New York 3, New York.

Copyright 1957 by THE AMERICAN SOCIOLOGICAL SOCIETY.

Entered at the Post Office at Baltimore, Maryland, as second-class matter.

SOCIOMETRY

A JOURNAL OF RESEARCH IN SOCIAL PSYCHOLOGY

VOLUME 20 • NUMBER 4 • DECEMBER 1957

CONTENTS

The Diffusion of an Innovation Among Physicians
James Coleman, Elihu Katz and Herbert Menzel 253

Brainwashing, Conditioning, and *DDD* (Debility, Dependency, and Dread) ..... *I. E. Farber, Harry F. Harlow and Louis Jolyon West* 271

Role-Taking Accuracy and Adjustment ..... *Sheldon Stryker* 286

Group Performance Under Feedback that Confounds Responses of Group Members ..... *Robert L. Hall* 297

Intercity Telephone and Airline Traffic Related to Distance and the "Propensity to Interact" ..... *Carl Hammer and Fred Charles Iklé* 306

Cumulative Scaling as a Routine Procedure with the IBM 101
Edgar F. Borgatta 317

SOCIOMETRY

EDITORIAL POLICY

Sociometry is concerned with the entire range of interests and problems represented by research in social psychology. It is the policy of the editors to seek those manuscripts for publication which represent the significant research interests of investigators who are concerned with giving the field of social psychology theoretical structure and reporting research which is clearly focused, well designed, and competently conducted.

While social psychology is presently regarded by most as a field with indeterminate boundaries, it has as its central focus the investigation of the processes and products of social interaction at the interpersonal, intrapersonal, intergroup and intragroup levels and the development of significant generalizations therefrom. In keeping with the more general meaning of the name of the journal emphasis will be placed on measurement of social behavior. However, this emphasis does not exclude the acceptability of good articles which must rely upon qualitative materials and analyses.

The editors and editorial consultants can be expected to subject manuscripts to rigorous criticism and screening according to the best standards of scientific research and at the same time avoid a sterile orthodoxy which would stultify the communication of creative ideas at the growing edge of the science. Thus the journal will strive to be flexible in its response to the publication needs of its contributors.

It is the intention of the editors to avoid any tendency toward professional provincialism and to invite contributions from any sector of the scientific community which promise to further the objectives of the journal.

Annual *subscription* rate $9.00; special rate to members of THE AMERICAN SOCIOLOGICAL SOCIETY $4.50. Single issues $2.50.

Specifications for Manuscript Preparation for Sociometry may be obtained from the Editorial Office.

Back Issues dated 1937 through 1955 should be ordered from the former publisher, Beacon House, Beacon, New York.

Notice of *change of subscriber's address* should be sent to THE SOCIETY four weeks in advance, indicating the old as well as the new address.

The Diffusion of an Innovation among Physicians[1]

JAMES COLEMAN, ELIHU KATZ, *University of Chicago*

HERBERT MENZEL, *Columbia University*

Anthropologists and sociologists have long been concerned with the processes through which customs, practices, attitudes, or messages spread. Traditionally, these processes have been studied by examining the ecological distribution of the trait at successive points in time. In a few cases, the actual transmission of messages from person to person has been traced out (e.g., 1, 3, 4, 5, 10). A still different approach to the study of this problem is reported in this paper. The population is physicians in four cities; the item whose use was spreading was a new drug; and the study focused on the ongoing social processes which finally led to widespread adoption of the drug by these physicians.

Data were collected 15 months after a new drug with wide potential use, here called "gammanym," had been placed on the market. By this time almost all the doctors in relevant specialties in the four cities studied had used the drug, some almost immediately, others only after a considerable interval of time. The research problem, stated most concretely, is this: What were the social processes which intervened between the initial trials of the drug by a few local innovators and its final use by virtually the whole medical community? The results reported below concern the effectiveness of networks of interpersonal relations at each stage of the diffusion process. The study is to be reported in full elsewhere (2); a pilot study has already been reported upon (9). A separate article by one of us describes the cumulative research experiences which led to the decision to focus explicitly upon interpersonal relations, using sociometric techniques (6).

METHODS—I

The method of survey research, involving structured interviews with a sample of physicians, was used. But since the problem as defined

[1] This article may be identified as Publication No. A 239 of the Bureau of Applied Social Research, Columbia University. An earlier version was read at the annual meeting of the American Sociological Society, Detroit, Michigan, September 8, 1956. We are indebted to Helmut Guttenberg for creative assistance throughout the project. Philip Ennis, Marjorie Fiske, Rolf Meyersohn, and Joseph A. Precker participated in the design of this study. The preparation of this paper was facilitated by funds obtained from a grant made to the Bureau of Applied Social Research by the Eda K. Loeb Fund.

concerned the social structure which linked these doctors together, it was necessary to deviate in two important ways from the customary survey design which, in effect, treats individuals as so many independent units of observation. (a) Each doctor interviewed was asked three sociometric questions: To whom did he most often turn for advice and information? With whom did he most often discuss his cases in the course of an ordinary week? Who were the friends, among his colleagues, whom he saw most often socially? In response to each of these questions,[1] the names of three doctors were requested. This made it possible to trace out the links by which each doctor was connected with the rest of the medical community. (b) It was decided to include in the sample, as nearly as possible, *all* the local doctors in whose specialties the new drug was of major potential significance. This assured that the "others" named by each doctor in answer to the sociometric questions were included in the sample, so that it became possible to characterize pairs or chains of socially connected doctors. Accordingly, 125 general practitioners, internists, and pediatricians were interviewed; they constituted 85 per cent of the doctors practicing in these fields in four Midwestern cities, ranging in population from 30,000 to 110,000.[2]

The dependent variable of the analysis which follows is the month during which each doctor first used the drug. This information was *not* obtained in the interviews; it was obtained through a search of the prescription records of the local pharmacies for three-day sampling periods at approximately monthly intervals over the 15 months following the release date of gammanym. In this way, the month during which each doctor first used the drug was ascertained.[3] The research is thus based on three kinds of data: the month of each doctor's first prescription for the new drug, obtained through a search of pharmacists' files; data about the informal social structure of the medical community, derived from doctors' replies to sociometric questions in an interview; and many individual attributes of each doctor, likewise obtained by interview.

[2] In addition, 103 doctors in other specialties were also interviewed, thus making a total sample of 228, or 64 per cent of all doctors in active private practice in these cities. The analysis presented here is based only on the 125 general practitioners, internists, and pediatricians, except that sociometric designations accorded them by the remaining 103 doctors were included when measuring the sociometric status of the 125.

[3] The date so ascertained will tend to be slightly later than the doctor's actual introduction date, due to the sampling of days. The interval between sampling periods was made to alternate between 32 and 25 days, so that each two successive sampling periods included all 6 days of the working week. Records were obtained from 64 of the 84 drug stores in the four cities. Of the remaining 20, only two had any significant pharmaceutical business.

RESULTS—I

Before presenting the results concerning interpersonal relations, the results concerning other ("individual") determinants will be briefly characterized. As expected, the date on which a doctor first prescribed the new drug was related to a large number of his *individual* attributes, e.g., his age, the number of medical journals he subscribed to, his attachments to medical institutions outside his community, and certain attitudinal characteristics. To illustrate the relationship of drug introduction date to such individual attributes, one of the latter will be singled out: the doctor's relative orientation to his professional colleagues and to patients, inferred from his answer to the following question:

How would you rank the importance of these characteristics in recognizing a good doctor in a town like this?

a. The respect in which he is held by his own patients
b. His general standing in the community
c. The recognition given him by his local colleagues
d. The research and publications he has to his credit

The following rankings were classified as "profession-oriented": cdab, cadb, cbda, cabd; the following rankings were classified as "patient-oriented": abcd, acbd, acdb, bacd. The 14 doctors who gave other rankings were assigned to one group or another by a rank-order scaling procedure which will be described in detail elsewhere (2).

Figure 1 shows the relationship of the resulting classification to the date of introduction of the new drug. The solid curve represents those doctors who were classified as profession-oriented, and shows the cumulative proportion of gammanym users among them for each month. Thus, for example, by the fourth month 40 per cent of these doctors had used gammanym; by the sixth month over 50 per cent. The lower curve similarly represents the doctors who were classified as patient-oriented; by the sixth month only 42 per cent had used the drug. Thus the more profession-oriented doctors in these cities generally used the drug earlier than the less profession-oriented ones.[4] Similar results were obtained for many other individual attributes—i.e., attributes describing individuals without reference to their social relations with one another.

[4] The difference between the mean adoption dates of the two groups in Fig. 1 is 2.8 months, which is significant at the .01 level, using a standard two-tailed test of difference between means of normally distributed variables. It should be pointed out, however, that the argument of this report does not rest on the statistical significance of isolated findings so much as on the consistency of the results of several diverse approaches with one another and with prior theoretical notions. It is doubtful that significance tests in the usual sense are meaningful in situations like the present. For a detailed statement of our position in this matter, see (8, p. 427).

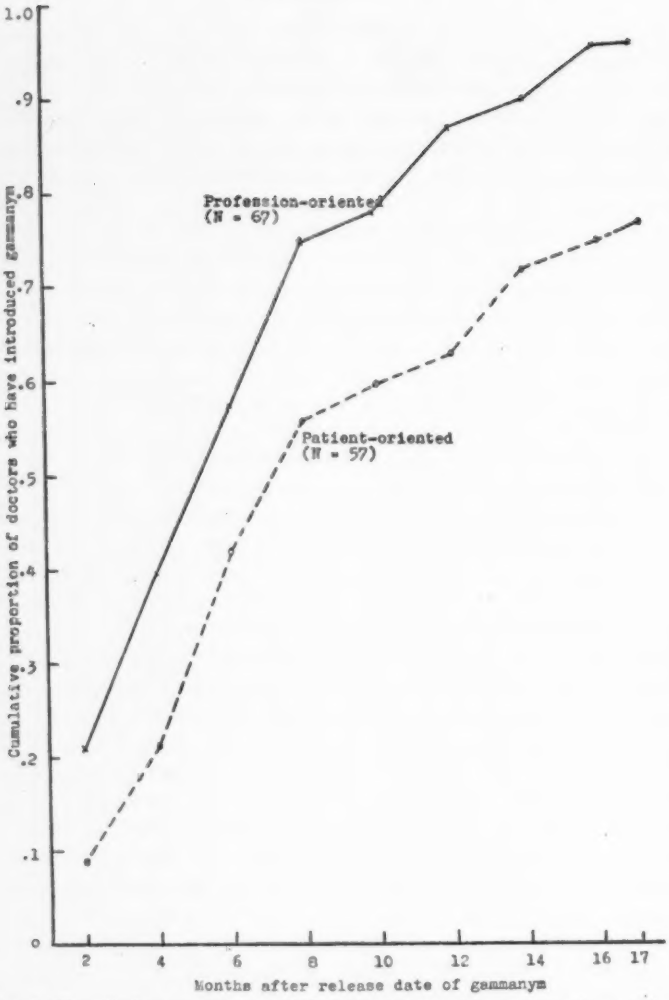


FIG. 1. Cumulative proportion of doctors introducing gammanym: profession-oriented *vs* patient-oriented.

But even stronger relations were found when we turned to *social* attributes—those characterizing a doctor's ties to his local colleagues. Doctors who were mentioned by many of their colleagues in answer to any of the three sociometric questions used the drug, on the average, earlier than those who were named by few or none of their colleagues. More generally

speaking, the degree of a doctor's integration among his local colleagues was strongly and positively related to the date of his first use of the new drug. Figure 2 shows, for example, the results with regard to the network of friendships. The "integrated" doctors—those named as "friends" by three or more of their colleagues—were much faster to introduce gammanym

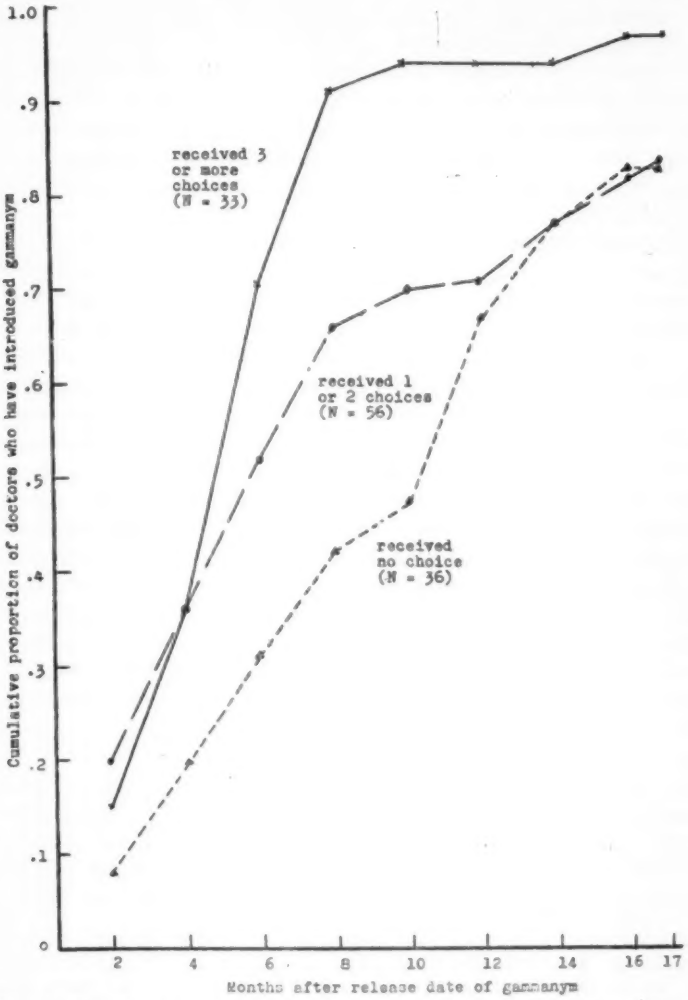


FIG. 2. Cumulative proportion of doctors introducing gammanym: differences in integration on friendship criterion.

into their practices than the rest. The networks of discussion and of advisorship yielded similar findings.

Two important contrasts differentiate Figure 2 from Figure 1, and, more generally, social attributes from individual ones, in their relation to gammanym introduction. First, the relationship in Figure 2 (as measured, for example, by the difference between the mean drug introduction dates of the extreme groups) is greater than that in Figure 1; greater, in fact, than the relationship of the introduction date of gammanym to all but one of the many individual characteristics which were examined. (The single exception is the doctor's total prescription volume for the general class of drugs which includes gammanym: the greater his use of drugs of this type, the earlier did he introduce gammanym.)[5] This emphasizes the importance of social contacts among doctors as a crucial determinant of their early use of the new drug.

But it may reasonably be questioned whether the relationship shown in Figure 2 may not arise merely because the measures of social integration are themselves associated with some personality or other individual differences which predispose a doctor to early introduction. It is in answer to this question that a second contrast between Figures 1 and 2 is relevant.

Notice that the two curves in Figure 1 are roughly parallel, differing from one another only in vertical displacement. This is true as well in most of the remaining charts (not shown) which relate individual characteristics to gammanym introduction. The curves in Figure 2, by contrast, differ from each other in shape as well as location: the curve for the more integrated doctors, although not starting out much higher than the other curves, rises steeply upward with a slight gain in slope at the fourth month, while the curve for the more isolated doctors rises at a moderate and almost constant slope. To put it differently, the integrated doctors were little different from their isolated colleagues at the very beginning; but then their rate accelerated to produce an increasing gap between the curves. In contrast, the profession-oriented doctors in Figure 1 differed from the patient-oriented from the very start almost as much as later on.

The constant difference between the profession-oriented and patient-oriented doctors suggests that they differ individually in their receptivity to new developments in medicine. On the other hand, the accelerating difference between the integrated and isolated doctors suggests a kind of "snowball" or "chain-reaction" process for the integrated: They are

[5] The difference between the mean drug introduction dates of those high and low on integration according to the 3 sociometric questions used is 3.1, 4.1, and 4.3 months. The difference between those with high and low total prescription volume for this general class of drugs is 5.0 months. Only one other individual characteristic (number of journals read) produced a mean difference of as much as 4.0 months.

individually little different in receptivity from their more isolated colleagues, but as their fellows come to use the drug, they pick it up from these doctors themselves; and as more of their fellows come to use it, their chances of picking it up are greater.

The difference between the two kinds of relationship to drug introduction is also shown by Table 1, which compares the individual variables and the social variables in their relation to gammanym introduction at two points in time: 1 month and 7 months after the drug was introduced. For each of these dates, the table shows the average difference in per cent of gammanym users (a) between those measuring "high" and "low" on each of twelve individual variables and (b) between those measuring "high" and "low" on three measures of social integration. The latter are based on choices received in response to the three sociometric questions mentioned earlier. The twelve individual variables include all those examined which showed a difference of two or more months in mean date of introduction between the high and the low groups.

The size of these differences measures the size of the relationship at the two times. As is evident, the social integration measures show a slightly *smaller* relationship than do the individual variables after 1 month, but a much *larger* relationship after 7 months. Thus, as exemplified by the comparison between Figures 1 and 2, the socially integrated doctors "pull away" from their isolated colleagues, while the doctors differing in some individual attribute simply maintain their intrinsically different receptivity as time goes on.

Figures 3 and 4 show the difference between two corresponding theoretical "models" of the introduction process. In Figure 3, the upper and lower curves both express a model of "individual innovation"; the difference between the two is simply that the receptivity is greater for the upper. This difference in individual innovation rate or receptivity corresponds, we suggest, to the difference between profession-oriented and patient-oriented doctors (and between doctors who differ in other individual attributes as

TABLE 1

The Average Relation of Twelve "Individual" Variables and of Three Measures of Social Integration to the Rate of Gammanym Introduction at Two Points in Time

	Average Difference in Per Cent of Gammanym Users between High and Low Groups		Ratio of Differences
	After 1 Month	After 7 Months	
Individual variables	9.2	27.4	2.98
Social integration	8.7	40.3	4.64

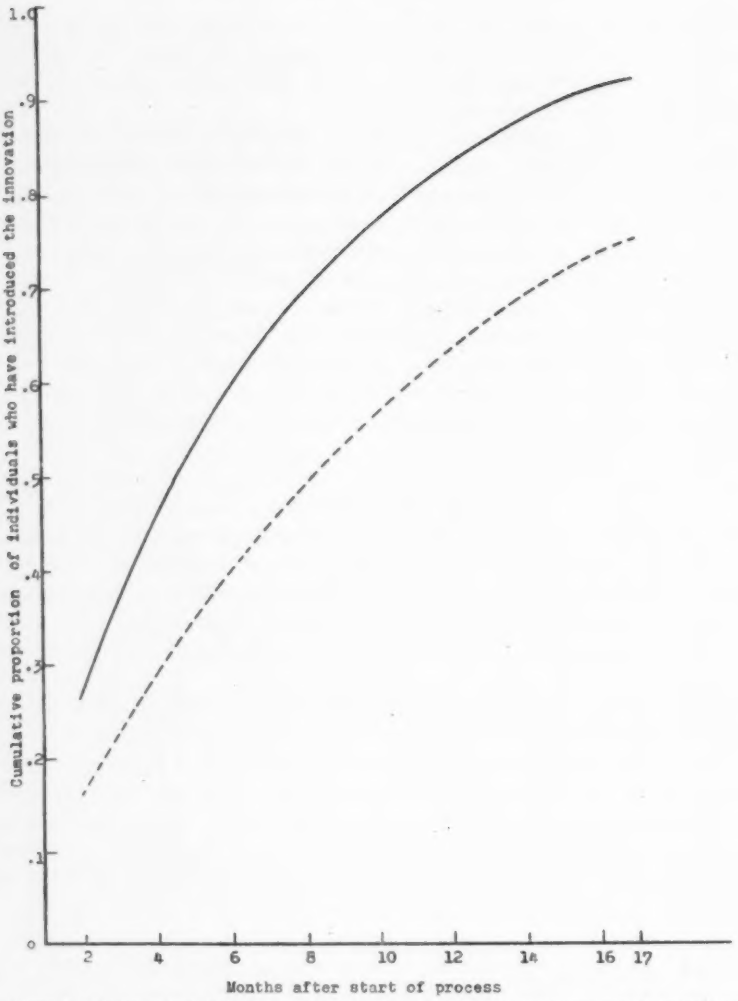


FIG. 3. Model of individual innovation, showing effects of differences in individual receptivity, k.

$$\frac{dy}{dt} = k(1 - y)t.$$

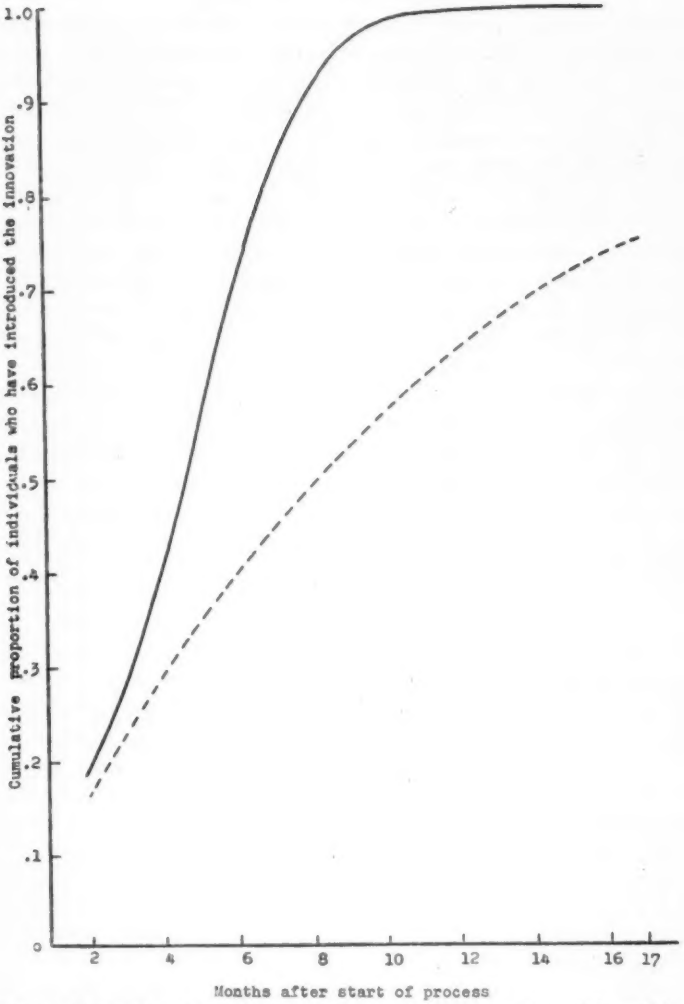


FIG. 4. Comparison of model of "chain-reaction" innovation with model of individual innovation.

$$\frac{dy}{dt} = ky(1 - y)t.$$

well). In contrast, in Figure 4 the upper curve (which is roughly similar in shape to the curve for the integrated doctors) represents a snowball process in which those who have introduced pass on the innovation to their colleagues. (This curve is described by an equation which has been used to characterize rates of population growth, certain chemical reactions, and other phenomena which obey a chain-reaction process.) The lower curve in Figure 4 is still the individual innovation process. (Technically, the individual and snowball processes are described by equations on the graphs, which can be paraphrased as follows: *Individual process*—the number of doctors introducing the new drug each month would remain a constant percentage of those who have not already adopted the drug. *Snowball process*—the number of doctors introducing the new drug each month would increase in proportion to those who have already been converted.)

In short, these comparisons suggest that the process of introduction for those doctors who were deeply embedded in their professional community was in fact different from the process for those who were relatively isolated from it. The highly integrated doctors seem to have learned from *one another*, while the less integrated ones, it seems, had each to learn afresh from the journals, the detail man (drug salesman), and other media of information.

METHODS—II

This result called for a more detailed investigation into the ways in which the networks of relations among the doctors affected their introduction of the new drug. Such an investigation required a shift of focus from doctors to relationships among doctors or to the networks themselves as the units of analysis. Various methods could have been devised to do this. We chose to record the behavior of *pairs* of doctors who were sociometrically related to one another, reasoning that if the networks of relations were effective, then pairs of doctors who were in contact must have been more *alike* in their behavior than pairs assorted at random. That is, if there was a snowball or chain-reaction process of drug introduction from one doctor to another, then adjacent links in the chain—pairs of socially related doctors—should have introduced the drug about the same time.

In order to test this hypothesis for the discussion network, Figure 5 was constructed. (Similar figures were constructed for the networks of friendship and advisorship.) Each sociometric pair was assigned to a column of this matrix according to the gammanym introduction date of the chooser, and to a row according to the gammanym introduction date of the doctor chosen. (A mutual choice constitutes two pairs in this tabulation, since any chooser and his choice constitute a pair.) Pairs of doctors who introduced the drug during the same month (interval zero) fall

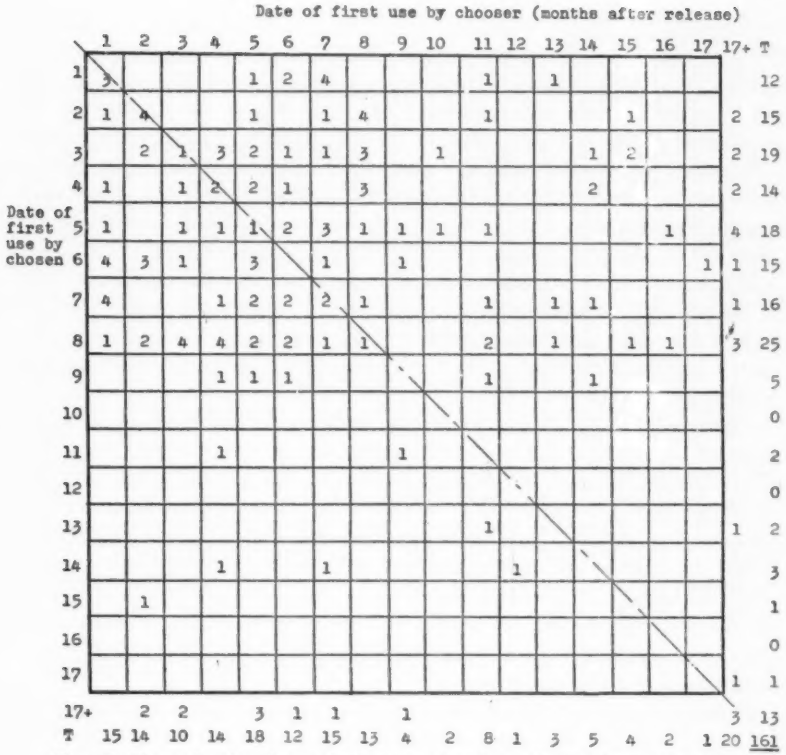


FIG. 5. Chart showing dates of adoption of each member of discussion pairs.

in the main diagonal; pairs of doctors who differed in introducing the drug by an interval of one month fall into cells adjoining the diagonal; and so on.

The resulting distribution of these intervals for the sociometric pairs was then compared to the corresponding distribution of intervals for a set of "random pairs" which has the following characteristics. If a pair is selected at random: (a) the probability that the chooser-member of the pair introduced gammanym during a particular month is the same as in the actual sample but is independent of the introduction date of the doctor chosen; (b) the probability that the chosen member introduced gammanym during a particular month is the same as in the actual sample but is independent of the introduction date of the doctor making the choice. Thus, for example, among the random pairs, those who introduced gammanym in the first month and those who did so in the seventh gave equal portions of their choices to other first-month introducers. Similarly, those who introduced

gammanym in the first month and those who introduced it in the seventh *received* equal portions of their choices from first-month introducers. Operationally, a set of "chance" frequencies satisfying these criteria can easily be obtained by computing for each cell of Figure 5 the product of the associated marginal totals, divided, for convenience, by the grand total.[6]

Contrary to expectations, the proportion of pairs whose members had introduced gammanym during the same month, one month apart, two months apart, and so on, according to the chance model proved to be almost identical to the proportion of actual discussion pairs who had introduced gammanym simultaneously or with varying intervals. The results for pairs of friends and for advisor-advisee pairs were similarly disappointing. This meant the rejection of our original hypothesis that pairs of doctors in contact would introduce the drug more nearly simultaneously than pairs of doctors assorted at random.

There was, on the other hand, the earlier evidence that the doctor's integration was important to his introduction of gammanym. This dictated a more intensive look at the behavior of pairs of doctors. Accordingly, we raised the question whether the networks, though ineffective for the *whole* period studied, may have been effective for the *early* period, immediately after the drug was marketed. An inspection of Figure 5 suggests that this could easily be the case. If only the upper left-hand portion of the matrix, representing the first two, three, or four months, is considered, then there appears to be a tendency for both members of a pair to introduce the drug in the same month.

In order to describe this tendency more precisely, it was decided to eliminate from consideration those associates of each doctor who used the drug only after *he* did. That is to say, the following question was now asked of the data: How closely did the drug introduction of each doctor follow upon the drug introductions of those of his associates who had introduced the drug before him? The answer is: very closely, for early introducers of the drug; not at all closely, for late introducers of the drug.

This result is based on a measure for each month, obtained by dividing up the total matrix of pairs of doctors as shown in Figure 6. The single cell in the upper left-hand corner represents those pairs both of whose members introduced the drug in the first month. The L-shaped section next to it contains the pairs which consist of one doctor who introduced the drug in

[6] A complication arose from the fact that the study was carried on in four different cities, with sociometric choices between cities excluded. This could spuriously raise any measure of pair-wise similarity of behavior, if there are large differences in behavior between the cities. (This fact was called to our attention by Jack Feldman of NORC). In order to avoid such a spurious relation, "chance" frequencies, as above described, were calculated separately from the marginal totals for each city, and only then summed over the cities.

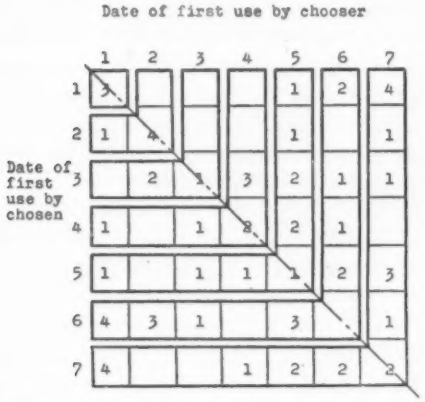


FIG. 6. Exploded view of portion of Fig. 5, showing monthly segments.

the second month and one who introduced it in the first *or* second. The next L-shaped section contains all pairs which consist of one third-month adopter and one third-month-or-earlier adopter, and so on. It was now possible to determine the average interval for the sociometric pairs in each L-shaped section; likewise the average interval for the corresponding random pairs. On this basis, a measure of simultaneity was computed for each section, according to the formula:

Measure of Simultaneity (positive) =

$$\frac{\text{(avge. interval for random pairs)} - \text{(avge. interval for sociometric pairs)}}{\text{avge. interval for random pairs}}.$$

This measure expresses the difference between the random and actual intervals as a fraction of the difference between the random interval and complete simultaneity (i.e., an interval of zero). The measure thus has a maximum of 1, and is zero when pairs are no closer than chance. In those cases where the actual interval exceeded the random interval, a different denominator was used.[7]

[7] Measure of simultaneity (negative) =

$$\frac{\text{(avge. interval for random pairs)} - \text{(avge. interval for sociometric pairs)}}{(s-1) - \text{(avge. interval for random pairs)}}$$

s being defined as the number of the latest month included in the particular L-shaped section. (E.g., $s = 4$ in the case of pairs consisting of one fourth-month adopter and one fourth-month-or-earlier adopter.) When the index has a negative value, it therefore expresses the difference between the random and actual intervals as a fraction of the difference between the random interval and the maximum interval that is possible.

RESULTS—II

The values of the index are plotted in Figure 7 for the second through the sixth months. Separate curves are plotted for pairs of friends, discussion pairs, and advisor-advisee pairs. The interpretation of these results must be tentative because of the small numbers of cases; on the other hand, the patterns which emerge are rather consistent.

Figure 7 suggests, first of all, that the networks of doctor-to-doctor contacts operated most powerfully during the first 5 months after the release of the new drug: such influence as any doctor's drug introduction had upon his immediate associates evidently occurred soon after the drug became available. (Figure 7 omits the later months during which the index is negative or very small.) Second, the three networks did not behave identically.[8] The discussion network and the advisor network showed most pair-simultaneity at the very beginning and then progressively declined. The friendship network shows initially less pair-simultaneity than the other two, but—with some instability—appears to reach its maximum effectiveness later. Finally, after the fifth or sixth month following the release of the new drug, none of the networks any longer showed pair-simultaneity beyond chance.

These results, however tentative, suggest that there may be successive stages in the diffusion of this innovation through the community of doctors. The first networks to be operative as chains of influence appear to be those which connect the doctors in the professional relationships of advisors and discussion partners. Only then, it seems, does the friendship network become operative—among those doctors who are influenced in their decisions more by the colleagues they meet as friends than by those whom they look to as advisors or engage in discussion during working hours. Finally, for those doctors who have not yet introduced the drug by about 6 months after the drug's release these networks seem completely *inoperative* as chains of influence. The social structure seems to have exhausted its effect; those doctors who have not responded to its influence by this time are apparently unresponsive to it. When they finally use gammanym, they presumably do so in response to influences outside the social network, such as detail men, ads, journal articles, and so on, and not in response to their relations with other doctors.

But one further phase in the social diffusion of gammanym can be discerned by examining separately the sociometrically integrated and the

[8] Many of the sociometric ties reappear in two or three of the networks. The three sociometric questions yielded a total of 958 "pairs" within the sample of 125 doctors; but since some of these pairs were identical in answer to two or all three of the questions, there were only 704 *different* pairs. This overlap is still small enough to allow differences in patterns to emerge, as shown in the text.

relatively isolated doctors. One would expect the networks of doctor-to-doctor contact to show their effectiveness first among the more integrated doctors and only then among those who are less integrated in their medical community. It has already been seen (Fig. 2 and text) that the more isolated doctors, on the average, introduced gammanym considerably later than the socially more integrated doctors. We now propose, however, that when more isolated doctors *did* introduce the drug early, it was not with the help of the social networks. While the networks were operative as channels of influence *early* for the integrated doctors, they were operative

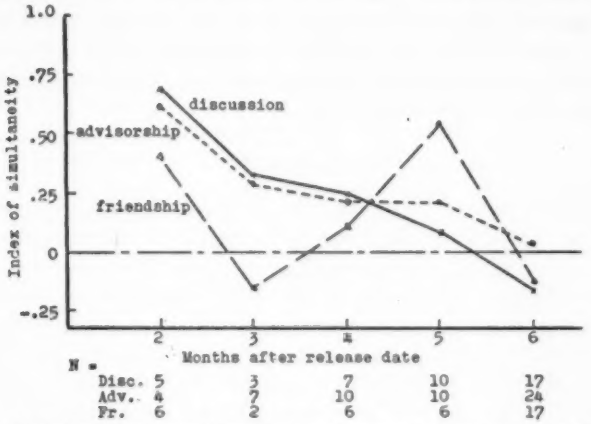


FIG. 7. Index of pair-simultaneity for three networks at different times.

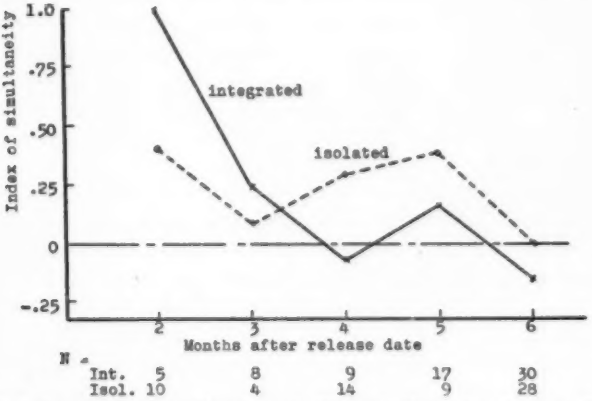


FIG. 8. Index of pair-simultaneity at different times for doctors differing in integration.

only later for the more isolated ones. This is what seems to have occurred. Figure 8 plots the index of simultaneity separately for more and less integrated doctors. (The graphs show weighted averages for all three networks; separately the numbers of cases would be so small as to produce erratic trends.)

The peak of effectiveness of doctor-to-doctor contacts for the well-integrated doctors appeared in the earliest month for which it can be plotted—the second month—after which effectiveness sharply declined. For the relatively isolated doctors, by contrast, the networks were not so effective at first as were those for the integrated doctors, but they maintained their effectiveness longer. Thus it appears that the networks of relations were effective not only for the more integrated doctors but also for the relatively isolated doctors who introduced the drug during the first 5 months of the drug's availability.

CONCLUSION

The above results, taken together, suggest a process which may be summarized as follows: At first the influence of these social networks operated only among the doctors who were integrated into the community of their colleagues through ties of a professional nature—as advisors or as discussion partners. Then it spread through the friendship network to doctors who were closely tied to the medical community through their friendship relations. By this time, social influence had also become operative in the more "open" parts of the social structure—i.e., among the relatively isolated doctors. Finally, there came a phase during which most of the remaining doctors introduced gammanym but did so in complete independence of the time at which their associates had introduced it: the networks now showed no effect. For the integrated doctors, this phase began about 4 months after the drug's release; for the isolated doctors, it began about 6 months after the drug's release. This picture is of course a tentative one, for the small size of the sample introduces variability, and there may be factors which produce spurious results.

There remains the question: Why should these sociometric ties to colleagues who have used the drug be influential during the first months of the drug's availability, but not later? One possible answer lies in the greater uncertainty about the drug that must have prevailed when it was new. (Data not reported here show that those doctors who introduced gammanym early did so far more tentatively than those who introduced it later.) We know from work in the tradition of Sherif that it is precisely in situations which are objectively unclear that social validation of judgments becomes most important.

More generally, this explanation implies that a doctor will be influenced more by what his colleagues say and do in uncertain situations, whenever and wherever they may occur, than in clear-cut situations. This explanation was confirmed by further data from the study which show that doctors influence each other more in treatments whose effects are unclear than in treatments whose effects are clear-cut. This topic will be dealt with in detail elsewhere (7).

CONCLUDING METHODOLOGICAL NOTE

A word should be added about the significance of research of this kind, aside from the possible interest in its specific substantive findings. It exemplifies a methodological approach which will, we feel, assume a larger role in the social research of the next decade: namely, making social relationships and social structures the units of statistical analysis. To be sure, the analysis of social relations has always been the sociologist's business. Nevertheless, most empirical studies have either treated and described a community, a factory, a hospital ward, or any other large grouping of people as a single unit, or else they have statistically analyzed data collected on hundreds or thousands of single inividuals, as in the typical "survey" study. What has been missing until recently is study designs which would explicitly take into account the structuring of single persons into larger units, and yet allow sophisticated quantitative treatment. The techniques of sociometry can meet this purpose, but have, with some notable exceptions (e.g., 4, 11), been applied chiefly to small closed groups and primarily for descriptive purposes.

The attempt reported here has been to carry out a design and analysis which would effect a marriage between sociometric techniques and survey research, in order to investigate quantitatively problems of the sort which community studies have ordinarily investigated by qualitative means. The attempt, of course, points up many more problems than it even partially solves: e.g., how to integrate an analysis of formal social structures with an analysis of informal ones; how to proceed from pair-analysis to the analysis of longer chains and complex networks; and so on. A set of methodological and substantive problems awaits the researcher. It is suggested that the solution will give sociologists important new tools with which to investigate social dynamics.

Manuscript received: April 4, 1957
Revised manuscript received: June 11, 1957
Herbert Menzel
Bureau of Applied Social Research
Columbia University
New York 25, N. Y.

REFERENCES

1. Back, K., *et al.*, "A Method of Studying Rumor Transmission," in L. Festinger *et al.*, *Theory and Experiment in Social Communication*, Ann Arbor, Michigan: Research Center for Group Dynamics, University of Michigan, 1950, 307–312.
2. Coleman, J. S., E. Katz, and H. Menzel, "The Diffusion of a Medical Innovation" (tentative title), Glencoe, Illinois: The Free Press (in preparation).
3. Dodd, S. C., "Diffusion is Predictable," *American Sociological Review*, 1955, 20, 392–401.
4. Festinger, L., S. Schachter, and K. Back, *Social Pressures in Informal Groups*, New York: Harper, 1950, Chap. 7.
5. Jennings, H. H., "Leadership and Isolation," in G. E. Swanson, T. M. Newcomb, and E. L. Hartley, eds., *Readings in Social Psychology*, New York: Holt, 1952.
6. Katz, E., "The Two-Step Flow of Communication: An Up-to-Date Report on an Hypothesis," *Public Opinion Quarterly*, 1957, 21, 61–78.
7. Katz, E., J. S. Coleman, and H. Menzel, "Social Influence on Decision-Making in Ambiguous Situations—Some Data from a Survey among Physicians" (unpublished).
8. Lipset, S. M., M. A. Trow, and J. S. Coleman, *Union Democracy*, Glencoe, Illinois: The Free Press, 1956.
9. Menzel, H., and E. Katz, "Social Relations and Innovation in the Medical Profession: The Epidemiology of a New Drug," *Public Opinion Quarterly*, 1956, 19, 337–352.
10. Moreno, J. L., *Who Shall Survive?*, Beacon, New York: Beacon House, 1953, 440–450.
11. Riley, M. W., J. W. Riley, Jr., and J. Toby, *Sociological Studies in Scale Analysis*, New Brunswick, New Jersey: Rutgers University Press, 1954.

Brainwashing, Conditioning, and *DDD* (Debility, Dependency, and Dread)[1]

I. E. FARBER, *State University of Iowa*

HARRY F. HARLOW, *University of Wisconsin*

AND

LOUIS JOLYON WEST, *University of Oklahoma School of Medicine*

Few aspects of Communism have been more puzzling and disturbing to the Western world than the widely publicized collaboration, conversion, and self-denunciation in individuals—communist and noncommunist, innocent and guilty alike—who have suffered Communist imprisonment. Such behavior in persons whose intelligence, integrity, or patriotism can scarcely be doubted has suggested to many a mysterious power or knowledge that enables Communists to manipulate the thoughts and actions of others in a manner ordinarily reserved to characters in the more lurid sorts of science fiction. Accordingly, such terms as "brainwashing," "thought control," "menticide," and so on, have been applied to the process or product of this manipulation. To lend some degree of scientific respectability to such concepts, attempts have been made (e.g., 12, 16) to relate them to the psychiatric implications of Pavlovian conditioning procedures.

While these speculations have an undeniable romantic appeal, more sober analyses (1, 2, 7) of factors influencing the behavior of prisoners under Communist control indicate that they are neither mysterious nor indicative of any unusual amount of psychiatric sophistication on the part of Communists. Indeed, considering the extraordinary degree of control the Communists maintain over the physical and social environments of their prisoners, it is rather surprising that their efforts to indoctrinate and convert have not been more successful. Contrary to the views of some writers in popular media, the record indicates that most American prisoners in Korea, for instance, showed remarkable "sales resistance," even under profound duress.

It is a fact that the Communist Chinese in Korea achieved considerable

[1] This paper is a revision and elaboration of a report for the Study Group on Survival Training, sponsored by the Air Force Personnel and Training Research Center, March, 1956, of which the authors were members. The initial report was prepared by the first two authors, and the research for the revision was supported in part by the United States Air Force under Contract No. AF 41(657)-75 monitored by the Director, Officer Education Research Laboratory, Maxwell Air Force Base, Alabama, with the third author as Chief Investigator. Permission is granted for reproduction, translation, publication, use, and disposal in whole and in part by or for the United States Government.

success in stimulating cooperative behavior in a large number of United Nations prisoners of war through a combination of threats, propaganda, group pressures, and group manipulation. By Segal's criteria, 15 per cent of American army prisoners cooperated unduly. And if it can be considered that it was every man's duty to exercise active resistance to the enemy and his propaganda during the period of captivity, then fully 95 per cent failed to meet the most stringent criteria for commendable behavior (18, 28, p. 80). Nevertheless, the Chinese induced only 21 American prisoners to remain under Communism (13), and it is doubtful whether all these were truly "converted." Most authorities agree that despite occasional lapses the vast majority of American prisoners of war performed well and honorably. As the Secretary of Defense's advisory committee on POW's has reported, "the record seems fine indeed" (30).

In the light of these findings, a complete analysis would concentrate more heavily on the factors that enabled the large majority of POW's to resist in some degree. However, it is not with these phenomena that the present discussion is primarily concerned. Rather, we wish to discuss the basis for the success of techniques whereby false confessions, self-denunciations, and participation in propaganda activities were brought about. The Communists made special efforts to elicit these behaviors in flying personnel, particularly with regard to confessions of participation in bacteriological warfare. After their world-wide propaganda campaign went into high gear with accusations of "germ warfare" in Korea, beginning on February 21, 1952, a vigorous policy of coercive pressure was applied to a large number of American flying personnel captured during the Korean conflict. As a result, a number of flyers from the Air Force and Marine Corps signed false confessions of bacteriological warfare and participated to various extents in enemy propaganda activities. A detailed account of these events may be found elsewhere (24, 28).

The objective intensity of noxious stimulation, injury, disease, malnutrition, deprivation, sleeplessness, fatigue, isolation, and threat suffered by many prisoners for a greater or lesser period was extreme. There were few, if any, who were not subjected to some of these conditions. Accounts of observations and experiments related to these various types of stress are now appearing in the literature in increasing numbers (e.g., 11, 29). The present discussion is concerned with the theoretical analysis of the psychological states and processes resulting from such objective conditions of stress.

DDD

Although the specific components of these states vary in intensity and pattern, in the case of the prisoner of war they contain at least three important elements: debility, dependency, and dread. They refer to the fact

that individuals subjected to the kinds of environmental conditions listed above have reduced viability, are helplessly dependent on their captors for the satisfaction of many basic needs, and experience the emotional and motivational reactions of intense fear and anxiety. These components are separable, but it is evident that they also interact. Consequently it seems appropriate as well as convenient to conceive of these states and processes as though they were an entity or syndrome including debility, dependency, and dread, to be referred to as *DDD*. Among the POW's pressured by the Chinese Communists, the *DDD* syndrome in its full-blown form constituted a state of discomfort that was well-nigh intolerable.

Debility was induced by semi-starvation, fatigue, and disease. Chronic physical pain was a common feature. Loss of energy and inability to resist minor abuse, combined with the lack of proper facilities for the maintenance of personal hygiene, led to inanition and a sense of terrible weariness and weakness.

Dependency, produced by the prolonged deprivation of many of the factors, such as sleep and food, needed to maintain sanity and life itself, was made more poignant by occasional unpredictable brief respites, reminding the prisoner that it was possible for the captor to relieve the misery if he wished. If an individual was placed in prolonged isolation, as was so often the case with flyers pressed to confess to the bacteriological warfare charges, the deprivation of ordinary social stimulation and relations markedly strengthened the dependency. Although we shall not dwell on this aspect of the situation, the effectiveness of Communist methods was undoubtedly greatly enhanced by their control of the means for satisfying nuclear social needs for recognition, status, communication, and so on. The captors' condemnation and misunderstanding of American social values, in connection with the withdrawal of accustomed social supports, e.g., reliable sources of information and communication with others as a means of testing reality and of appraising moral standards, played a significant part in the dependency relationship (2, 7, 10, 17).

Dread is the most expressive term to indicate the chronic fear the Communists attempted to induce. Fear of death, fear of pain, fear of nonrepatriation, fear of deformity or permanent disability through neglect or inadequate medical treatment, fear of Communist violence against loved ones at home, and even fear of one's own inability to satisfy the demands of insatiable interrogators—these and many other nagging despairs constituted the final component of the *DDD* syndrome (2).

The interrelations of these factors, carefully contrived and nurtured by the Communists, were of great importance in determining the total effect of *DDD*. Although there were some individuals who acceded to the demands of their captors fairly early in the game, it is clear that the Chinese realized the importance of preparing the resistant prisoner, through *DDD*, for the

long, drawn-out process designed to bring about the desired goal—complete compliance.

Before considering in greater detail the specific mechanisms underlying the role of *DDD* in accomplishing this aim, three prefatory comments are in order. First, the present analysis lays no claim to comprehensiveness. It deals with only a few aspects of *DDD* occurring under certain conditions. We believe these aspects to be important, but they are not all that is important. In this connection, the present paper may be considered as an elaboration of portions of the comprehensive discussion of Communist "thought reform" by Hinkle and Wolff (7). It is gratifying that our conclusions, arrived at independently and on somewhat more theoretical grounds, are essentially in agreement with theirs.

Second, our use of the terminology of learning theory, broadly conceived, and our use of concepts derived from conditioning, does not imply that we consider learning theory uniquely competent to explain the effects of *DDD*. On the other hand, we do consider factors influencing behavior in *DDD* to have something in common with factors affecting behavior in learning situations generally, and, therefore, that it may be worth while attempting to analyze some aspects of behavior associated with *DDD* in terms of principles of classical and instrumental conditioning. But, as an eminent conditioning theorist has recently noted (20), the view that principles derived from conditioning might apply to more complex behavior does not at all imply that complex behavior can be explained solely in terms of the variables affecting conditioning. In this instance, it is particularly doubtful that the procedures used to influence the behavior of prisoners under Communism derived from the methods of Pavlov, or that the prisoners' reactions are generally understandable in purely Pavlovian terms. On the contrary, to the extent that such concepts apply at all, selective or instrumental (Thorndikean) learning was a more prominent feature than classical (Pavlovian) conditioning. Certainly, only limited aspects of the behavior of prisoners under Communism bear any resemblance to the generalized inhibitory or excitatory states characterizing some of Pavlov's dogs (14).

Finally, we should beware of the "psychologist's error." Although some of the behavior of prisoners under Communism may be susceptible to analysis in terms of learning and conditioning principles, it does not follow that the application of these principles by Communist captors was deliberate and self-conscious. Animal trainers and side-show barkers are often extremely competent manipulators of behavior; this does not mean they are comparative or social psychologists.

DDD, SELF-PERCEPTION, AND THINKING

By providing a radically changed context *DDD* might be expected to produce new responses that actively compete or interfere with wonted

behavior. It may also produce a condition of markedly reduced responsiveness, not unlike the generalized inhibitory states described by Pavlov (14) and Liddell (8), due to the reduced or monotonous stimulation associated with isolation and confinement, or to reduced energy, or to the frustration of previously successful techniques for achieving goals. Whenever individuals show extremely selective responsiveness to only a few situational elements, or become generally unresponsive, there is a disruption of the orderliness, i.e., sequence and arrangement of experienced events, the process underlying time spanning and long-term perspective. By disorganizing the perception of those experiential continuities constituting the self-concept and impoverishing the basis for judging self-consistency, *DDD* affects one's habitual ways of looking at and dealing with oneself.

This effect, which has elsewhere been related to the collapse of certain ego functions (22), bears an interesting resemblance to some aspects of the postlobotomy syndrome. The latter, too, is characterized by apathy and the disturbance of the self-concept or self-regarding tendency (15). The frequency and degree of flattened affect and self-deprecation in the confessions of prisoners under Communism have probably been overestimated, but to the extent they have occurred, the observed behavior has much in common with that of some brain-damaged individuals.

Closely related to the foregoing consequence of *DDD* is a disturbance of association and a concreteness of thinking similar to that sometimes seen in schizophrenia. The retention of recent experiences and habit patterns may be impaired, with consequent regression, i.e., primitivization, in language, thought, and those integretative and mediating symbolic processes essential to reasoning and foresight. Conditioning performance in human subjects is impaired by some kinds of symbolic activity, and conversely, the impoverishment of thinking may increase susceptibility to arbitrary and unsubtle training procedures (cf. 3) leading to relatively automatic and uncritical imitative responses. This susceptibility may be further enhanced by anxiety and emotionality (5, 21, 23).[2]

[2] These assumptions do not imply a negative correlation between intelligence and conditioning in normal subjects, nor better conditioning in feeble-minded or brain-damaged subjects than in normal individuals. The empirical evidence does not support any such views. The suggested effect of impoverished thinking relates only to that produced by debility, isolation, and such factors. One may speculate, in this connection, on the relation between this putative effect of *DDD* and the kinds of hypersuggestibility and automatism reported among primitive peoples suffering from prolonged physical stress and privation. Whether these symptoms result from some state of hyperconditionability is a moot question. Arctic hysteria and latah, for instance, are presumably dissociative and therefore hysteroid in nature (25), and the relation between hysteria and conditioning is as yet uncertain (4, 5, 6). Thus it is not possible at present to identify the effects of *DDD* with any particular psychiatric state.

REINFORCEMENT OF SOCIAL COMMUNICATION

On the assumption that conditioning principles apply in part to the behavior of prisoners of war, it is important to analyze further the nature of the conditioned stimuli and the responses elicited by them. Careful consideration would seem to indicate that the situation contains features both of selective or instrumental learning and of classical conditioning (20). The instrumental (i.e., Thorndikean rather than Pavlovian) aspect is emphasized by the fact that an individual must acquire a particular set of responses in order to bring about a reinforcing state of affairs. It is our thesis that an alleviation in the state of *DDD* provides the reinforcement for much of the behavior desired by the enemy. In other words, *DDD* does not, in and of itself, produce the desired behavior. *DDD* merely provides the occasion for the selective reinforcement of certain modes of response.

The role of *DDD* in the reinforcement process depends on the fact that it is not constant. Instead, it may be assumed to fluctuate in time, partly as a result of spontaneous psychophysiological processes, and partly as a result of deliberate manipulations designed to maintain its intermittent nature (2), thus preventing its fall to a baseline of permanent depression and hopelessness. Those individuals who were reduced to complete apathy undoubtedly represented failures from the point of view of their Communist captors.

At the risk of considerable oversimplification, one may conceive of two consequences of the occasional mitigation of *DDD*. First is the conditioning of the "expectancy" that *DDD* will be alleviated. (This constitutes the actual classically conditioned anticipatory goal response.) Relief, whether due to spontaneous factors or deliberate manipulations, is intermittent, temporary, and unpredictable. Far from weakening the expectancy of relief, however, this tends to maintain the expectancy and renders it less susceptible to extinction. In nontechnical terms, this process serves to keep hope alive, permitting some degree of adaptive behavior, and inhibiting self-destructive tendencies, which would frustrate the enemy's purpose.

This aspect of the learning process throws some light on the frequent practice in Communist prisons of having prisoners "punish themselves." Thus, a captive might be instructed to stand or kneel in a certain position until he should decide to cooperate. This emphasis on the self-inflicted nature of the prisoner's punishment, and his ability to mitigate his condition "voluntarily," is clearly calculated to increase the intensity of expectancies of the possibility of relief. At the same time, it is evident that the prisoner's belief that he actually exercises control is delusory, so far

as the objective facts are concerned, since the captor may select any behavior he chooses as the condition for relieving a prisoner's distress.

The alleviation of *DDD* at the time of occurrence of the desired behavior leads to the second consequence—the learning of instrumental acts. This is not so difficult to arrange as one might suppose and is certainly not the result of any mysterious power of the manipulator. Very often, the desired behavior is verbal in nature. Verbal behavior is in a general way already strongly conditioned to *DDD* in all human adults. One learns from infancy to use verbal behavior as a means of relieving or avoiding many of the components of *DDD*. And, as the foregoing discussion indicates, the aperiodic and unpredictable nature of the selective reward of particular language responses may be one of its chief strengths. If one may extrapolate from the results of numerous laboratory experiments, this is the very procedure calculated to produce the maximum number of responses and also to make them highly resistant to extinction, even in the absence of rewards (19).

The nature of the rewards used needs no elaboration. Relief of hunger, fatigue, isolation, or pain, even temporarily, serves as an automatic reward. Even the verbal and empty promise of alleviation of *DDD* leads to appropriate anticipatory goal responses, keeping hope alive. Paradoxically, interrogation, harangues, threats, and contumely may also have a rewarding aspect, so great is the acquired reinforcement value of social communication and speech under conditions of isolation, dependency, and physical debility.

Since the habits of social communication associated with *DDD* are initially strong, and are further strengthened by selective reinforcement, it is not strange that prisoners often show considerable social responsiveness in the presence of their captors. Despite the impoverishment of the self-concept and primitivization of thinking referred to earlier, prisoners could enjoy in some degree a much needed social relationship in the interrogation and indoctrination situations. It may be hypothesized that some prisoners became the victims of the very socialization process that under ordinary circumstances is regarded as a desirable and, indeed, essential aspect of civilized living. It is of interest in this connection to record the finding of Lifton, who explicitly noted among a group of repatriated prisoners who had most aggressively resisted collaboration with the Communists, a large portion of individuals with significant antisocial tendencies (9). We do not suggest that collaboration and confession by prisoners under Communism are signs of desirable social attitudes. We do suggest that socialization training facilitates the tendency to engage in social communication, even with a recognized enemy, particularly under conditions in which the behavior is reinforced by the satisfaction of power-

ful drives while at the same time interfering or inhibitory tendencies are markedly reduced.

There are some analogies between the condition of an individual under such circumstances and that of a hypnotized subject. The hypnotized subject also tends to respond automatically, especially to verbal stimuli, to be greatly influenced by the attitude of the hypnotist, and to be highly selective in his social responsiveness. Furthermore, there is general agreement regarding the susceptibility of most normal individuals to hypnosis, except in the case of strong deliberate resistance. Under conditions of *DDD*, the possibility of resistance over a very long period may be vanishingly small. As soon as resistance appears, the intensity of *DDD* can be increased, thus at one and the same time punishing resistance and increasing the influence of the reward when relief occurs. It must be remembered that the strengthening effects of rewards—in this instance the alleviation of an intensely unpleasant emotional state—are fundamentally automatic. They occur because of the kind of nervous system we have, and not in any essential way because of the mediation of conscious thought processes.

RETENTION OF PRISON EXPERIENCES AND BEHAVIOR

What is the aftermath of such experiences? The evidence clearly indicates that, except in the case of organic brain damage such as might result from avitaminosis, the behavior of the typical returnee from Communist prisons is "normal," in the special and important sense that he behaves in a manner that would be predicted on the basis of ordinary laws of behavior. There is not the slightest evidence for the necessity of postulating new or unknown factors or conditions. This does not mean the experience of imprisonment leaves no trace. Such a circumstance would in itself be abnormal, i.e., inconsistent with the known principles of behavior. In terms of normative criteria, many ex-prisoners are more than ordinarily anxious, defensive, dependent, suspicious, insecure. Pressed to explain any possibly discreditable acts, they often exhibit a very considerable degree of hesitancy, vagueness, paramnesia, and rationalization. In a word, they behave exactly as one would expect of any individual required to explain and defend his behavior, many determinants of which he is not aware.

Most returnees remember a great deal of what occurred during their imprisonment. They do not remember everything and may be unable to give a very clear account of their own behavior. Some behavior may appear as strange and inexplicable to the person concerned as to anyone else. The explanation of whatever impairment of memory occurs may be found in the laws of forgetting, deriving from both clinic and laboratory. There is no need to expatiate here on the role of repression in forgetting when the material to be recalled elicits anxiety and guilt. But it may be useful to

note briefly some of the factors that would influence retention even in the absence of these emotions.

In an earlier section, it was pointed out that the state of *DDD* produces responses that actively compete with ordinary responses to environmental stimuli. By the same process, the comforting and familiar stimuli of home and friends are associated with a wholly different set of responses from those produced by *DDD*. The changed context may actively inhibit recall of the prison experiences. This phenomenon is nothing more than the familiar psychological explanation of forgetting in terms of associative interference.

Among the most important of these competing responses are the affective ones. The returnee simply does not feel as he did as a prisoner. He may be able to talk about how he felt, although this too offers difficulties because our terminology for describing emotional states is woefully inadequate and vague (3), but he does not currently respond affectively in the same way. Similarly, the familiar stimuli of home reinstate different verbal responses, both overt and implicit, that affect recall. The returnee feels different, talks differently, and thinks differently than he did in the former context. Since, like all of us, he is unaware of many of the cues to his former behavior (as well as his current behavior), it is as useless to ask him to explain his earlier reactions as it is to ask a person why he once disliked olives or is for the moment unable to recall the name of an old acquaintance.

The particular reactions and attitudes constituting patriotism, bravery, loyalty, and so on, depend on the appearance of particular cues, symbolic or other. Such qualities are tendencies to respond positively or negatively, in varying degrees and combinations, in the presence of certain combinations of cues. From this point of view, unwonted reactions occurring under *DDD* do not represent a different attitude; rather, the habitual attitude does not appear because the appropriate cues have been removed. Back home in the presence of adequate cues, the returnee tends to act and feel as he did prior to imprisonment.

Finally, one must consider the effect on retention of the adequacy of the original impression. Occasionally the returnee does not remember much because he did not observe much. The impoverished stimulation, impaired responsiveness, reduced symbolic activity, and disorganization of time-spanning characteristic of *DDD* reduce the clarity and strength of impressions at the time of the original experience, and thus decrease ability to recall.

In the light of all these factors, whose pejorative influence on retention is well known by students of human learning, it is clearly to be expected that the recall of returnees would be something less than complete and wholly accurate as regards their actual prison experiences and behavior.

RESISTANCE TO EFFECTS OF *DDD*

Despite our opinion that the most undesirable effects of *DDD* are not necessarily permanent, or, given appropriate social conditions after repatriation, even particularly long-lived, the general picture of *DDD* presented in the foregoing discussion is rather gloomy. This is in part because we have emphasized its stressful aspects rather than the considerable resources most persons can muster to oppose them. The many environmental, social, and motivational variables that produce resistance to these effects have not been discussed, but their potency should certainly not be underestimated. As we have observed earlier, the resistance of American prisoners under Communism, in the face of the objective circumstances detailed above, was in most instances notable, and in some nothing less than heroic (2, 24, 28).

It is evident that there are great individual differences in susceptibility to *DDD* even under conditions in which the level of *DDD* itself could reasonably be regarded as constant, i.e., not a differential factor. To state the point somewhat differently, there are unquestionably a number of variables, whose values differ from person to person, affecting the degree of resistance to the effects of *DDD*. The question may then be raised whether the potency of these variables might not be increased in any given individual. We believe they can.

The statement, "Every man has his breaking point," contains a germ of truth, but like other bromides, is liable to misinterpretation. It does not mean the "breaking point" is fixed for any given individual, so that nothing can affect it. Such a view is scientifically indefensible, if not meaningless, since it implies that some kinds of behavior are unlawful, i.e., not affected by variations in any kinds of antecedent conditions. Furthermore, the term "breaking point" is itself misleading. Susceptibility to *DDD* or any other stressful condition is not an all-or-none affair. We are discussing behavior, and behavior varies in degree and in kind. It may be possible to define "breaking" in the manner that one defines a right or wrong response in arithmetic, but it should be recognized that such a definition would be arbitrary at best and of doubtful conceptual significance. As Biderman has pointed out, a prisoner's physical and moral strength may be sapped by Communist coercive methods to a degree that resistance appears insignificant. But, however feeble his performance, motivation to resist usually persists and shows itself as circumstances permit (2).

It is not the purpose of the present discussion to consider all the possible personal or social variables of which resistance to the effects of *DDD* may be a function, or indeed to consider any of them in detail. We mention two, not because they are necessarily of particular importance, but because they throw further light on the nature of the *DDD* state. First, there is

the factor of physical health. Other things equal, there is probably a negative relation between degree of physical health and vigor on the one hand and susceptibility to *DDD* on the other. Debility can be postponed longer, dependency fought against, and the self-concept maintained more easily if bodily well-being obtains. Second, there is the factor of initial or chronic anxiety. No matter what anxiety is due to, the higher the anxiety level, the greater is the possibility of rewarding behavior by its momentary reduction. Contrariwise, a low level of initial anxiety should retard the growth of the "dread" component of *DDD*, and at least indirectly affect some of its antecedents, e.g., the reactivity to pain (27).

Thus, techniques for promoting health and decreasing anxiety in those who may become prisoners are probably of great importance. Nevertheless, one should not expect factors such as these to block the effects of *DDD* indefinitely. Physical health, for instance, may be of utmost value over the short haul, e.g., during early interrogation. But on a long-term basis it may be relatively insignificant. Health can be broken down by a determined and informed enemy in a very short time. And although a healthy individual can better resist the effects of debilitating variables, there is no evidence that, once illness and physical debility occur, previously healthy individuals can tolerate this condition better than those who might have become habituated to it. In some cases, indeed, the reverse might obtain.

A somewhat similar reservation may be expressed concerning procedures calculated to reduce initial anxiety, i.e., training individuals to be generally nonanxious. The fear component of *DDD*, unlike neurotic anxiety or neurotic fears (phobias), is quite realistic for the most part. Realistic fears are not easily extinguishable and, if they were, the desirability of extinguishing such fears is not altogether certain. For instance, fear of punishment for displaying hostility toward one's captors is adaptive. Wolf and Ripley (26) quote one prisoner of the Japanese in World War II in this regard: "I had to make a conscious effort not to resent things because I realized that my bones are brittle."

On the other hand, certain anticipatory fears may be modified through training procedures. Alleviation of unrealistic fears of the unknown (through accurate indoctrination regarding enemy methods) undoubtedly improves the ability of the individual to deal with those fears that are realistic. It may make it possible for him to admit his fear to himself, as a reasonable and expected reaction, thus modifying its influence as a covert force toward compliance. Furthermore, an expectation of the probable psychophysiological effects of stress may rob them of some of their "shock" value. Finally, a certain amount of transfer may be expected from stressful training experiences in which adaptive modalities have been learned, thus permitting the prisoner to conceptualize his current stressful experience

in terms of previous (and at least partly successful) transactions under stress.

Still, it would be foolish to disregard the fact that some of the elements of *DDD* represent a pathological organic state, some consequences of which are probably innately determined. To the extent this is true, one cannot expect to achieve a great degree of prophylactic success in regard to the effects of *DDD*, any more than one can reasonably expect at the present state of knowledge to prevent some of the undesirable consequences of lobotomy.

Though many of the behavioral consequences of *DDD* are not innately determined, the conditioning of certain types of responses desired by the enemy may eventually occur, even in the face of superlative resistance. One of the conclusions that may legitimately be drawn from the present analysis of the circumstances of imprisonment under Communism is that, if a prisoner's state of *DDD* reaches a truly extreme degree of severity (and it cannot now be predicted whose ability to resist will be the most effective in combating *DDD*), and *if he lives*, he probably cannot be expected to resist indefinitely. This prediction does not require the assumption that Communists have mysterious powers, or that their prisoners are subjected to some strange process of "brainwashing" negating the effects of their previous training and attitudes. It is based, rather, on the assumption that under the physical, social, and emotional conditions of extreme *DDD*, some degree of ultimate compliance may be considered a natural consequence of the operation of ordinary principles of human behavior.

SUMMARY

Although the behavior of some prisoners under Communism, including collaboration, conversion, and self-denunciation, appears to suggest that Communists are able to "brainwash" their prisoners in a mysterious way, a consideration of the physical, emotional, and social conditions of the prisoner in conjunction with the ordinary principles of human behavior reveals that such behavior may be readily explained. The state of the prisoner may be described in terms of the concepts of debility, dependency, and dread (*DDD*), and some of the behavioral principles explaining the effects of the *DDD* state derive from learning and conditioning phenomena.

It is assumed that *DDD* operates in part to produce a generalized state of hyporesponsiveness, disrupting time-spanning processes and disorganizing the self-concept. Another consequence of *DDD* is the impairment of symbolic processes, perhaps rendering the prisoner susceptible to relatively simple conditioning techniques. The intermittent nature of *DDD* leads both to the expectancy of relief (i.e., hope) and to the reinforcement of

specific kinds of verbal behavior. The latter effect is facilitated by the fact that social communication is already strongly conditioned to cues such as those produced by *DDD*, as a result of normal socialization training.

The typical prisoner returnee exhibits no extraordinary peculiarities of memory. The degree of forgetting of prison experiences is such as would be expected as a result of the inhibition of anxiety-producing thoughts (repression), change of situational context during recall, and the inadequacies of original impressions during imprisonment.

Resistance to the undesirable consequences of *DDD* is a matter of degree and may be modified by such factors as physical health and level of initial anxiety. Nevertheless, factors such as these cannot reasonably be expected to provide more than temporary respite. Through various defenses, a prisoner may postpone the development of extreme *DDD* for a long time, perhaps indefinitely. But if a prisoner's state of *DDD* is extreme, and if he lives, he probably cannot resist indefinitely. Far from furnishing proof of the operation of some unnatural process of "brainwashing," this eventuality is a predictable consequence of the operation of laws of normal human behavior.

Manuscript received: May 17, 1957
Revised manuscript received: July 12, 1957
I. E. Farber
Department of Psychology
State University of Iowa
Iowa City, Iowa

REFERENCES

1. Bauer, R. A., "Brainwashing—Psychology or Demonology," American Psychological Association Symposium, September 3, 1956.
2. Biderman, A. D., "Communist Techniques of Coercive Interrogation," *Air Force Personnel and Training Research Center Development Report TN-56-132*, Lackland Air Force Base, Texas, 1956.
3. Dollard, J., and N. E. Miller, *Personality and psychotherapy*, New York: McGraw-Hill, 1950.
4. Eysenck, H. J., "Cortical Inhibition, Figural After Effect, and Theory of Personality," *Journal of Abnormal and Social Psychology*, 1955, 51, 94–106.
5. Farber, I. E., K. W. Spence, and H. P. Bechtoldt, "Emotionality, Introversion-Extraversion, and Conditioning," Midwestern Psychological Association, May 3, 1957.
6. Franks, C. M., "Conditioning and Personality: a Study of Normal and Neurotic Subjects," *Journal of Abnormal and Social Psychology*, 1956, 52, 143–150.
7. Hinkle, L. E., and H. G. Wolff, "Communist Interrogation and Indoctrination of 'Enemies of the State,'" *Archives of Neurology and Psychiatry*, 1956, 76, 115–174.
8. Liddell, H. S., "Conditioned Reflex Methods and Experimental Neurosis," in J. McV. Hunt (ed.), *Personality and the Behavior Disorders*, New York: Ronald, 1944.

9. Lifton, R. J., "Home by Ship: Reaction Patterns of American Prisoners of War Repatriated from North Korea," *American Journal of Psychiatry*, 1954, 110, 732–739.
10. Lifton, R. J., " 'Thought Reform' of Western Civilians in Chinese Communist Prisons," *Psychiatry*, 1956, 19, 173–196.
11. Lilly, J. C., "Effects of Physical Restraint and of Reduction of Ordinary Levels of Physical Stimuli on Intact, Healthy Persons," *Group for the Advancement of Psychiatry, Symposium No. 2, Illustrative Strategies for Research on Psychopathology in Mental Health*, New York: GAP Publication Office, 1790 Broadway.
12. Meerloo, J. A. M., "Pavlovian Strategy as a Weapon of Menticide," *American Journal of Psychiatry*, 1954, 110, 809–813.
13. Pasley, V., *21 Stayed*, New York: American Book—Stratford Press, 1955.
14. Pavlov, I. P., *Conditioned Reflexes and Psychiatry*, W. H. Gantt (trans.), New York: International, 1941.
15. Robinson, M. F., and W. Freeman, *Psychosurgery and the Self*, New York: Grune and Stratton, 1954.
16. Santucci, P. S., and G. Winokur, "Brainwashing as a Factor in Psychiatric Illness," *Archives of Neurology and Psychiatry*, 1955, 74, 11–16.
17. Schein, E. H., "The Chinese Indoctrination Program for Prisoners of War," *Psychiatry*, 1956, 19, 149–172.
18. Segal, J. "Factors Related to the Collaboration and Resistance Behavior of U. S. Army PW's in Korea," *Technical Report 33, Human Resources Research Office*, The George Washington University, Washington, D. C., 1956.
19. Skinner, B. F., *Behavior of Organisms*, New York: Appleton-Century, 1938.
20. Spence, K. W., *Behavior Theory and Conditioning*, New Haven, Conn.: Yale University Press, 1956.
21. Spence, K. W., and I. E. Farber, "Conditioning and Extinction as a Function of Anxiety," *Journal of Experimental Psychology*, 1953, 45, 116–119.
22. Strassman, H. D., M. B. Thaler, and E. H. Schein, "A Prisoner of War Syndrome: Apathy as a Reaction to Severe Stress," *American Journal of Psychiatry*, 1956, 112, 998–1003.
23. Taylor, J. A., "The Relationship of Anxiety to the Conditioned Eyelid Response," *Journal of Experimental Psychology*, 1951, 41, 81–92.
24. West, L. J., "U. S. Air Force Prisoners of the Chinese Communists," Group for the Advancement of Psychiatry Symposium on Prisoners of War, Nov. 11, 1956.
25. West, L. J., "Hypnosis and the Dissociative Reactions," *Journal of Clinical and Experimental Hypnosis* (in press).
26. Wolf, S., and H. S. Ripley, "Reactions among Allied Prisoners Subjected to 3 Years of Imprisonment and Torture by Japanese," *American Journal of Psychiatry*, 1947, 104, 180–193.
27. Wolff, H. G., and S. Wolf, *Pain* (3d ed., 1952), New York: Charles C. Thomas, 1952.
28. "Communist Interrogation, Indoctrination and Exploitation of American Military and Civilian Prisoners," *Hearings Before the Permanent Subcommittee on Investigations of the Committee on Government Operations, U. S. Senate, 84th Congress, 2nd Session*, Washington, D. C.: U. S. Government Printing Office, 1956.

29. "Factors Used to Increase the Susceptibility of Individuals to Forceful Indoctrination: Observations and Experiments," *Group for the Advancement of Psychiatry, Symposium No. 3,* New York: GAP Publications Office, 1790 Broadway.

30. "POW: The Fight Continues After the Battle," *Report of the Secretary of Defense's Advisory Committee on Prisoners of War,* Washington, D.C.: U. S. Government Printing Office, 1955.

Role-Taking Accuracy and Adjustment[1]

SHELDON STRYKER, *Indiana University*

The assumption that knowledge is necessarily adjustive and ignorance or lack of knowledge necessarily maladjustive may fairly be said to suffuse the thinking of current social science whether in its "pure" or "applied" forms (8, 19). This assumption is clearly, and perhaps especially, embedded in various therapeutic uses of role-playing techniques (2, 7, 20, 21, 27). It is present as well in the use of role-playing as a teaching aid (4, 15) and in the increased use of this technique in diverse business settings (3, 11, 17). In all these instances, it is assumed that role-playing will improve role-taking skills and that the increases in knowledge of others consequent to improved role-taking will mean more adjusted social relationships.

This line of thinking takes on wider interest with the observation that many who write in the tradition of George Herbert Mead argue the adjustment functions of accurate role-taking (5, 16, 22).[2] Here, it must be stressed that Mead's usage of the adjustment concept differs significantly from contemporary usage. For Mead, adjustment is synonymous with adaptation, both terms referring to a *process* whereby one alters the course of his behavior in terms of the demands of the social situation (18, pp. 155–159). Adjustment, as it is currently used, is a static concept, typically referring to an *end-state* of happiness, satisfaction, and the like (10, pp. 48–49). Again, however, the implication of adjustment in the sense of happiness and satisfaction as an end-product of accuracy in the role-taking process had been drawn by many.

Thus, on both practical and theoretical grounds, the answer to the question of the relationship between role-taking and adjustment (as end-state) becomes important. This question may be formulated as an hypothe-

[1] This research was partially supported by a grant from The Graduate School, Indiana University.

[2] While our interest in the question being raised stems from Mead, and thus the paper is oriented to his work and interpretations of it, the fact that persons of somewhat different theoretical leanings present analogous arguments should not be overlooked. So, for example, Sullivan (26) discusses "selective inattention," describing this process as a "security operation" by which "we fail to recognize the actual import of a good many things we see, hear, think, do and say, not because there is anything the matter with our zones of interaction with others but because the process of inferential analysis is opposed by the self system." Sullivan clearly regards selective inattention in negative terms, a "powerful brake on personal and on human progress" and "more than any other of the inappropriate and inadequate performances of life, the classic means by which we do not profit from experience . . ." (26, pp. 374, 346).

sis: the adjustment of the individual is a function of the accuracy with which he can take the role of the other(s) implicated with him in some social situation.

This hypothesis and the design of the study stem from an interest in testing implications drawn from Mead's theory. Given this interest, we have utilized a conception of role-taking as anticipation of the responses of others implicated with one in an ongoing social situation (18, pp. 151, 242–243, 253–254). Accurate role-taking has been operationally defined as the correct prediction of the responses of others. Role-taking, so conceived, is obviously related to such concepts as empathy, insight, social sensitivity, and so on, studies of which have used similar measurement procedures. These concepts are not, however, synonyms; and the similarity of measurement procedures should not be taken to imply identical conceptualization or theoretical concern.[3]

The role-taking conception requires that predicter and other be mutually implicated in an ongoing social situation, and that predictions be made with reference to responses meaningfully related to that social situation. These, and consideration of desirable controls, led to the selection of the family as the most appropriate setting for the study. Data pertinent to the hypothesis were gathered in the summer of 1954 from 46 family units,

[3] There is considerable confusion in the literature with regard to the concepts role-taking, empathy, sympathy, identification, insight, social sensitivity, etc. These are sometimes seen as independent, sometimes as overlapping, and sometimes as essentially identical. We cannot attempt, here, to distinguish systematically among these concepts: Dymond (9) provides an extensive review of this literature, offering necessary distinctions. Suffice it to say that, for us, role-taking is anticipatory behavior; and that the emotional unity, participation in emotional life of others, "feeling-with" quality, and fellow-feeling that sometimes (although not always) are incorporated into these other concepts are not seen as part of the role-taking concept.

It is true that there may be diverse sources of accuracy when role-taking scores are based on correctness in prediction. From the standpoint of Mead's theory, which postulates that one's behavior is predicated on the symbolic anticipation of responses of others, the "impurity" of our role-taking index is of little consequence. This theory, as it has thus far been developed, does not require distinguishing the various possible bases of accuracy. The problem of "impurities" in studies of empathy, insight, etc., has been attacked by Hastorf and Bender (13), Cronbach (6), and Gage and Cronbach (12). Hastorf and Bender suggest that "projection" can be screened out by subtracting from an accuracy score a second score based on the coincidence of ascriber's own responses and his predictions for another. This technique is unable, however, to distinguish between correct ascriptions based on knowledge and correct ascriptions based on projection when, in fact, ascriber and other respond identically to an item. The technique has the same difficulty with correct predictions when ascriber and other disagree with respect to an item (12). Cronbach (6) develops a model which distinguishes four components of accuracy scores. The model is, however, premised on a research design which requires judges to predict responses of (the same) multiple others, rather than on a design strictly analogous to ours.

residents of Bloomington, Indiana. Each family unit consisted of a married pair and the parents of one, but not both, of this pair.[4] Each subject responded to a 20-item Likert-type attitude scale dealing with traditionalism in family-related matters.[5] He was asked for his own attitudes, and to predict or ascribe the responses the two other-generation members of his family would make to the items of this scale. An index of role-taking accuracy, consisting of the number of correct predictions person A makes of person B's responses, expressed as a percentage of total predictions of B's responses, was computed for each subject.[6]

From the same set of responses, an index of agreement between parent and offspring was devised, consisting of the percentage of items on which a given parent's own responses to the items of the scale corresponded with his offspring's responses.

Subjects also responded to a 40-item instrument indexing their adjustment to each of the persons for whom they ascribed attitudinal responses, and to a 10-item instrument indexing their dependence upon these persons. Odd-even correlations of from .88 to .92, and a test-retest correlation of .87, deriving from various pretests of the adjustment index provide evidence that this index meets minimal reliability requirements. That the index is, to a degree, a valid instrument is indicated by its ability to distinguish between groups whose adjustment scores may be expected to differ: e.g., the scores of females with reference to their own relatives are higher than are those of males; scores with reference to own relatives are higher than those with reference to in-laws. We have no reliability data for the dependence index. Validity evidence includes the facts that, in accord with expectations, females indicate a greater degree of dependence upon their relatives than do males, and that there is greater "neutrality" on this index in relation to in-laws than in relation to own relatives. Further, unpublished data show that offspring scores indicating dependence on parent

[4] In most instances, only the parents of one of the offspring pair resided in the community. When both sets of parents were available, the choice of one was determined by the desire to maintain a balance in the numbers within relationship categories. It should be stressed that "sampling" was not random. We have used statistical techniques based on random sampling, nonetheless, on the grounds that formal (objective) assessment of our data was preferable to nonformal (subjective).

[5] This instrument was adapted from a scale devised by P. E. Huffman (14) and entitled "Traditional Family Ideology Scale." There is evidence from this source that the instrument is reasonably reliable and valid.

[6] Subjects responded to the attitude items via a four-point continuum: Strongly Agree, Agree, Disagree, Strongly Disagree. The decision to use this set of response categories was reached after extensive pretesting designed to provide items which were relatively stable both when a subject responded for himself and when he predicted how another would respond. The requirement of item stability also dictated that the two agree, and the two disagree, responses be scored as identical for purposes of the role-taking index.

TABLE 1

Parent Adjustment by High or Low Role-taking Accuracy

	F-S	F-D	F-IS	F-ID	M-S	M-D	M-IS	M-ID	Row
High									
M	56.70	60.10	54.23	48.67	59.58	58.42	50.58	58.64	55.71
N	10	10	13	12	12	12	12	11	92
Low									
M	59.15	64.15	53.70	58.45	61.55	66.72	55.45	57.08	59.66
N	13	13	10	11	11	11	11	12	92

Completed Analysis of Variance

	Sums of Squares	d.f.	Mean Square	F
Columns	2372.58	7	338.94	2.24
Rows	616.95	1	616.95	4.07
Interaction	627.82	7	89.69	
Within groups	25470.82	168	151.61	

Weighted mean difference, rows: 3.68. $F_{.05}$ (1,168) = 3.90.

are associated with parent scores indicating dominance over offspring, and that offspring scores indicating dominance over parent are associated with parent scores indicating dependence on offspring.[7]

The study has available, then, indices of role-taking accuracy, agreement, adjustment, dependence, and subjects' own attitudes on family traditionalism.

Each subject ascribed the attitudes of two others. There are, consequently, 16 categories of subjects—8 parent and 8 offspring—each containing the responses of 23 persons. The notations in subsequent tables follow the convention of placing the role-taker first, the person whose role is taken second, and use the following abbreviations: son—S; son-in-law—IS; daughter—D; daughter-in-law—ID; father—F; father-in-law—IF; mother—M; mother-in-law—IM. Thus, for example, the notation S-IF refers to the category comprised of male offspring who ascribed the attitudes of their fathers-in-law.

The expectation from the initial hypothesis is that high role-taking accuracy will be associated with high adjustment. To test this expectation, subjects in the various relationship categories were dichotomized, using median role-taking or ascription accuracy scores as the breaking point. The eight parent and eight offspring categories were grouped and analyzed separately, using analysis of variance.

Table 1 presents the test of the hypothesis for the parent categories.

[7] For more complete accounts of these indices, and the data on the basis of which many of these remarks are made, see (24). The unpublished data will be made available to anyone interested in them.

TABLE 2

Offspring Adjustment by High or Low Role-taking Accuracy

	S-F	S-M	S-IF	S-IM	D-F	D-M	D-IF	D-IM	Row*
High									
M	42.90	57.50	51.89	47.00	58.67	63.46	47.09	52.18	52.91
N	10	10	9	12	12	13	11	11	88
Low									
M	55.62	48.08	48.79	49.73	61.18	60.10	37.75	42.25	50.13
N	13	13	14	11	11	10	12	12	96

* Preliminary analysis of variance indicated no statistically significant difference between row means appears in this table.

The cell entries are mean adjustment scores, with higher scores indicating better adjustment. There is, it will be noted, a statistically significant difference between the weighted row means. However, the direction of the difference is not in accord with that hypothesized. That is, parents who are comparatively *poor* role-takers are significantly *better* adjusted with reference to their offspring than are parents who are comparatively accurate role-takers. The row mean difference in Table 2, for the offspring categories, is in the expected direction. However, this difference is not statistically significant.

What might explain the inverse, statistically significant relationship between role-taking accuracy and adjustment for the parent category?[8] Clues are provided through speculation concerning the parent-adult offspring relationship and the cited findings, and through a consideration of thinking and research already in the literature.

Theorizing about the parent-adult offspring relationship, the following assertions seem justified: (a) the self-identifications and respect of parents are bound up with the lives of their *adult* offspring to a greater degree than are those of such offspring with the lives of their parents. That is,

[8] It has been pointed out to me by Dr. Nathan L. Gage that the inverse relationship between role-taking accuracy and adjustment appearing in Table 1 could occur if there is also an inverse relationship between parent adjustment to offspring and offspring scores on the family traditionalism scale; and that, under these circumstances, the inferences drawn from the table are not warranted. Similarly, the implications drawn from subsequent Tables 3–8 would not be warranted if a negative relationship between parent adjustment and offspring traditionalism existed for the subjects represented in Tables 3, 5, and 7, and if a positive relationship between parent adjustment and offspring traditionalism obtained for the subjects represented in Tables 4, 6, and 8. These possibilities have been checked, using Pearsonian correlations and also using analysis of variance. The results of these analyses were negative; i.e., relationships between parent adjustment and offspring traditionalism which would call for an alternative interpretation of the findings did not occur.

(b) parents are psychologically more involved with their offspring than are offspring with their parents. Thus, (c) evidence of gaps between the attitudes and behaviors of parents and offspring are more likely to be resisted and distorted by parents; and (d) when such evidence becomes open and recognized by parents, its consequences for their adjustment to their offspring are likely to be accentuated.

More generally, Moore and Tumin argue that "... ignorance must be viewed not simply as a passive or dysfunctional condition, but as an active and often positive element in operating structures and relations" (19, p. 787). Robin Williams writes: "It is extremely important also that many of the implicit understandings that make society possible are not just *implicit* but are also resistant to statement: it is as if there is a tacit agreement not to express or become aware of what would be dysfunctional.... We suspect that a study of areas of blocked communication would often reveal conflicts that remain non-disabling only so long as they are kept from overt crystallization" (29, pp. 528–529). Pertinent are the researches of William F. Whyte (28) and Howard S. Becker (1). Whyte, describing the social structure of the restaurant, points out that, in a society in which males are expected to originate action for females, interposing a high counter as a physical barrier to free communication between waitresses and countermen filling their orders serves to protect the countermen. Becker suggests that platforms, pianos, chairs, and so on, perform the same function for the dance musician in relation to his audience.

The foregoing suggests an hypothesis which would serve to explain the

TABLE 3

Parent Adjustment, High-tradition Parents Only, by High or Low Role-taking Accuracy

	F-S	F-D	F-IS	F-ID	M-S	M-D	M-IS	M-ID	Row
High									
M	53.50	57.20	51.00	42.50	57.80	61.57	48.86	60.20	53.80
N	4	5	7	6	5	7	7	5	46
Low									
M	67.83	62.50	59.83	60.25	65.00	65.40	60.60	60.83	62.87
N	6	8	6	4	6	5	5	6	46

Completed Analysis of Variance

	Sums of Squares	d.f.	Mean Square	F
Columns	1622.09	7	231.73	2.13
Rows	1588.41	1	1588.41	14.60
Interaction	586.78	7	83.83	
Within groups	8270.83	76	108.83	

Weighted mean difference, rows: 8.42. F_{99} (1,76) = 6.99.

TABLE 4

Parent Adjustment, Low-tradition Parents Only, by High or Low Role-taking Accuracy

	F-S	F-D	F-IS	F-ID	M-S	M-D	M-IS	M-ID	Row
High									
M	58.00	66.60	60.50	54.83	63.50	64.60	61.50	61.20	60.98
N	7	5	4	6	4	5	4	5	40
Low									
M	51.50	63.20	48.33	57.43	57.13	59.00	46.57	51.14	54.12
N	6	5	6	7	8	6	7	7	52

Completed Analysis of Variance

	Sums of Squares	d.f.	Mean Square	F
Columns	1356.93	7	193.85	1.37
Rows	1012.96	1	1012.96	7.17
Interaction	587.69	7	83.96	
Within groups	10735.66	76	141.26	

Weighted mean difference, rows: 6.75. F_{99} (1,76) = 6.99.

finding that parents who are accurate in their role-taking are significantly less well adjusted to their offspring than are parents who are inaccurate role-takers, while the same finding does not hold for offspring role-takers. The hypothesis asserts that the greater one's vulnerability in a relationship, the greater the tendency to erect blocks to full knowledge of the other in that relationship.

Data are available to test this explanatory hypothesis. Earlier it was noted that the study provides scores on scales measuring traditional family attitudes, dependence, and agreement of parent and offspring with respect to familial attitudes. It may be expected that highly traditional parents will be more vulnerable in their relations to offspring than nontraditional parents; that parents who are dependent upon their children will be more vulnerable than independent parents; and that parents whose views do not agree with their offspring's will be more vulnerable than will parents whose views coincide with those of their offspring. Treating vulnerability as an intervening variable, from the hypothesis, then, it follows that the inverse relationship between role-taking and adjustment should be more marked for highly traditional parents than nontraditional, for dependent parents than independent, for low-agreement parents than high-agreement.

These expectations are examined in Tables 3–8. In each instance, analysis proceeded as follows: the parent sample was dichotomized, using median scores (on tradition, dependence, agreement). Then, each subsample was further broken down into two categories, accurate and inaccurate role-

takers, on the basis of median scores within these subsamples. The cell entries in the tables are, as earlier, mean adjustment scores.

As the analyses of variance indicate, the expectations deriving from the vulnerability hypothesis are borne out. For highly traditional parents (Table 3), dependent parents (Table 5), and low-agreement parents (Table 7), accurate role-taking is accompanied by comparatively poor adjustment. In each of these cases, row differences are statistically significant. For nontraditional parents (Table 4), independent parents (Table 6), and high-agreement parents (Table 8) row differences are *reversed*—i.e., the mean adjustment scores for accurate role-takers are higher than they are for inaccurate role-takers, although only in the case of the nontraditional

TABLE 5

Parent Adjustment, Dependent Parents Only, by High or Low Role-taking Accuracy

	F-S	F-D	F-IS	F-ID	M-S	M-D	M-IS	M-ID	Row
High									
M	52.75	57.00	58.00	48.33	57.57	64.43	59.86	60.57	57.65
N	8	4	6	6	6	7	7	7	51
Low									
M	64.80	66.29	57.17	64.17	67.40	67.83	63.00	62.33	64.00
N	5	7	6	6	5	6	7	6	48

Completed Analysis of Variance

	Sums of Squares	d.f.	Mean Square	F
Columns	1099.01	7	157.00	1.12
Rows	1119.86	1	1119.86	8.01
Interaction	642.53	7	91.79	
Within groups	11603.41	83	139.80	

Weighted mean difference, rows: 6.58. F_{99} (1,83) = 6.95.

TABLE 6

Parent Adjustmeni, Independent Parents Only, by High or Low Role-taking Accuracy

	F-S	F-D	F-IS	F-ID	M-S	M-D	M-IS	M-ID	Row*
High									
M	54.00	62.17	51.83	52.40	59.75	58.14	40.67	55.25	55.13
N	5	6	6	5	4	7	3	4	40
Low									
M	64.00	61.67	49.20	48.33	57.25	55.33	39.17	51.83	53.24
N	5	6	5	6	8	3	6	6	45

* Preliminary analysis of variance indicated no statistically significant difference between row means appears in this table.

TABLE 7

Parent Adjustment, Low-agreement Parents Only, by High or Low Role-taking Accuracy

	F-S	F-D	F-IS	F-ID	M-S	M-D	M-IS	M-ID	Row
High									
M	56.40	53.60	52.33	41.50	60.80	61.43	49.00	51.86	53.37
N	5	5	6	6	5	7	5	7	46
Low									
M	60.00	63.33	49.87	57.40	64.00	65.40	51.80	63.60	59.32
N	7	6	7	5	7	5	5	5	47

Completed Analysis of Variance

	Sums of Squares	d.f.	Mean Square	F
Columns	2484.87	7	354.98	2.52
Rows	793.71	1	793.71	5.63
Interaction	710.24	7	101.46	
Within groups	10848.37	77	140.89	

Weighted mean difference, rows: 5.89. F_{ss} (1,77) = 3.97.

TABLE 8

Parent Adjustment, High-agreement Parents Only, by High or Low Role-taking Accuracy

	F-S	F-D	F-IS	F-ID	M-S	M-D	M-IS	M-ID	Row*
High									
M	56.00	66.60	56.17	55.83	60.20	66.33	53.00	64.50	59.40
N	6	5	6	6	5	6	7	6	47
Low									
M	59.60	64.86	62.00	60.83	56.50	57.20	57.00	52.40	58.95
N	5	7	4	6	6	5	6	5	44

* Preliminary analysis of variance indicated no statistically significant differences between row means appears in this table.

parents is the difference significant, and in one case (Table 8) the row difference is extremely slight. It should be noted, however, that in the independent parent table (Table 6), in seven of the eight relationship categories there are higher mean adjustment scores in the accurate role-taker cells; one large "reversal" in the F-S category probably accounts for the absence of statistically significant results. There is, it may also be noted, no particular consistency in the direction of the row differences in the high-agreement parents table (Table 8).

In over-all terms, the evidence clearly supports the hypothesis that vulnerability increases the likelihood of blocks to communication.

To return to our point of departure, the presumed adjustive consequences of knowledge of others: our findings clearly call into question any easy assumption on this score. Rather, they point to the conclusion that, at least under certain circumstances, such knowledge is maladjustive. Further, these findings serve to emphasize the lack of clarity in interpretations of Mead's social psychological theory. It may still be true, as Mead held, that the process of adapting one's behavior to others with whom one is socially implicated is dependent on role-taking.[9] That one will always become better adjusted, in the sense of happier or more satisfied, through role-taking must, on the basis of these findings, be doubted.

The reverse side of this coin should be noted. The assertion that knowledge is necessarily adjustive cannot be defended; it is highly probable that the unqualified assertion that ignorance is adjustive is equally invalid. The problem, now, is to specify further the conditions under which knowledge or ignorance is adjustive or maladjustive.[10]

Manuscript received: May 3, 1957
Revised manuscript received: June 26, 1957
Sheldon Stryker
Department of Sociology
Indiana University
Bloomington, Indiana

REFERENCES

1. Becker, H. S., "The Professional Dance Musician and His Audience," *American Journal of Sociology*, 1951, 57, 136–144.
2. Borden, R., "The Use of Psychodrama in an Institution for Delinquent Girls," *Sociometry*, 1940, 3, 81–90.
3. Chapman, E. A., "Role-Playing in a Cooperative Retail Training Class," *Occupations*, 1951, 29, 358–359.
4. Coleman, W., "Role-Playing as an Instructional Aid," *Journal of Educational Psychology*, 1938, 39, 427–435.
5. Coutu, W., *Emergent Human Behavior: A Symbolic Field Interpretation*, New York: Knopf, 1949.
6. Cronbach, L. J., "Processes Affecting Scores on 'Understanding of Others' and 'Assumed Similarity,'" *Psychological Bulletin*, 1955, 52, 177–193.

[9] Recently, Steiner (23) has questioned the generality of this assertion, noting that empirical evidence with regard to the propositions that accurate social perceptions are responsible for interpersonal competence and for group efficiency is contradictory. His analysis of assumptions concerning collective activity leads him to conclude that these propositions will hold when "(a) group members are motivated to cooperate; (b) the accurately perceived qualities are relevant to the activities of the group; (c) members are free to alter their own behaviors in response to their perceptions of other members; and (d) the behavioral changes which are a consequence of accurate social perceptions are the kinds which produce a more thoroughly integrated dyadic system" (23, p. 273).

[10] For materials suggestive of further conditions, see (25).

7. Curran, F. J., "The Drama as a Therapeutic Measure in Adolescents," *American Journal of Orthopsychiatry*, 1939, 9, 215–232.
8. Davis, K., "The Application of Social Science to Personal Relations," *American Sociological Review*, 1936, 1, 236–247.
9. Dymond, R. F., *Empathic Ability: an Exploratory Study*. Ph.D. dissertation, Cornell University, 1949.
10. Foote, N. N., and L. S. Cottrell, Jr., *Identity and Interpersonal Competence*, Chicago: University of Chicago Press, 1955.
11. French, J. R. P., "Role-Playing as a Method of Training Foremen," *Group Psychotherapy*, New York: Beacon House, 1945.
12. Gage, N. L., and L. J. Cronbach, "Conceptual and Methodological Problems in Interpersonal Perception," *Psychological Review*, 1955, 62, 411–422.
13. Hastorf, A. H., and I. E. Bender, "A Caution Respecting Measurement of Empathic Ability," *Journal of Abnormal and Social Psychology*, 1952, 47, 574–576.
14. Huffman, P. E., "Authoritarian Personality and Family Ideology," M.A. thesis, Western Reserve University, 1950.
15. Kay, L. W., "Role-Playing as a Teaching Aid," *Sociometry*, 1946, 9, 263–274.
16. Lindesmith, A. R., and A. L. Strauss, *Social Psychology*, New York: Dryden, 1949.
17. Liveright, A. A., "Role-Playing in Leadership Training," *Personnel Journal*, 1951, 29, 412–416.
18. Mead, G. H., *Mind, Self, and Society*, Chicago: University of Chicago Press, 1934.
19. Moore, W. E., and M. M. Tumin, "Some Social Functions of Ignorance," *American Sociological Review*, 1949, 14, 787–795.
20. Moreno, J. L., "Psychodramatic Treatment of Marriage Problems," *Sociometry*, 1940, 3, 1–23.
21. Moreno, J. L., "Psychodramatic Treatment of Psychoses," *Sociometry*, 1940, 3, 115–132.
22. Newcomb, T. M., "Austistic Hostility and Social Reality," *Human Relations*, 1947, 1, 69–86.
23. Steiner, I. D., "Interpersonal Behavior as Influenced by Accuracy of Social Perception," *Psychological Review*, 1955, 62, 268–274.
24. Stryker, S., "The Adjustment of Married Offspring to Their Parents," *American Sociological Review*, 1955, 20, 149–154.
25. Stryker, S., "Relationships of Married Offspring and Parent: A Test of Mead's Theory," *American Journal of Sociology*, 1956, 62, 308–319.
26. Sullivan, H. S., *The Interpersonal Theory of Psychiatry*, New York: Norton, 1953.
27. Trendley, M. B., "Psychodrama and Social Case Work" *Sociometry*, 1944, 7, 169–177.
28. Whyte, W. F., "The Social Structure of a Restaurant," *American Journal of Sociology*, 1949, 54, 302–310.
29. Williams, R., *American Society: A Sociological Interpretation*, New York: Knopf, 1951.

Group Performance under Feedback that Confounds Responses of Group Members[1]

ROBERT L. HALL, *Lackland Air Force Base, Texas*

When a person working alone learns a task, he usually has available some direct knowledge of the effect of his actions upon the environment. On the other hand, when a person working in a group learns a task that requires more than one person, it is often true that the effect of his actions upon the environment is a net effect of the responses of several men. That is, the knowledge of results that is available to each group member is a single group output that combines in some fashion response measures of more than one individual. The *manner* in which the individual responses are combined into a single output measure may not be known to the group members, and a social psychologist studying the group may not even be able to specify how they are combined. In any case, each individual receives informative feedback that correlates imperfectly with his individual response. There is distortion in a group member's feedback, and the distortion should not be presumed simply to be random, either with respect to time or with respect to his own responses. The distortion is systematically related to the responses of his teammates and may in turn be related to his own responses which are generating a part of the feedback to his teammates.

Under these conditions, which we shall call conditions of "socially confounded feedback," there is no direct empirical evidence on the group performance curves and how they are affected by various training variables. The present paper reports an experimental study of the effects of two independent variables under conditions of socially confounded feedback in two-man groups. One independent variable is the type of pretraining, and the other is the relative weight accorded to the responses of the two teammates in the confounded feedback. In this study the combination of responses into a single output is a simple additive combination. There are two dependent variables: a measure of the accuracy of the group output and a measure of role differentiation (i.e., the extent to which members respond differently in contributing to a group output).

[1] This study was performed in support of Project 7713 under the Air Force Personnel and Training Research Center, Lackland Air Force Base, San Antonio, Texas. Permission is granted for reproduction, translation, publication, use, and disposal in whole and in part by or for the United States Government. This is a revised version of a paper read at Midwestern Psychological Association, Chicago, Illinois, May 3, 1957.

METHOD

Subjects: The subjects (Ss) were 64 male Air Force basic trainees, ranging in age from seventeen to about twenty-four, with mode at age seventeen. They were assigned as a unit for use as *Ss* in psychological research. *Ss* were divided into categories on the basis of their scores on the Armed Forces Qualification Test (AFQT). The categories, with their AFQT percentiles and roughly equivalent Stanford-Binet I.Q. (2) were:

I. 93 and above; I.Q. 127 and above
II. 65–91; I.Q. 108–126
III. 31–63; I.Q. 94–107

Those with AFQT scores lower than 31 were excluded. Subjects were used in pairs, randomly assembled with the restriction that both members of a pair be in the same AFQT category.

Apparatus: For each experimental session the two *Ss* were seated in identical adjacent booths, with walls about 5 feet high. They could not see each other, and verbal communication was not allowed. The booth wall in front of each *S* was a panel behind which a micrometer was mounted with its knob projecting about ¼ inch through the panel. A shield concealed the knob from the *S*, but he could easily reach under the shield to turn the knob. The experimenter (E), seated on the other side of the panel, could read the micrometer scale, which was a linear measure of the amount the knob was turned. On the panel in front of each *S* was a linear scale from 0 to 40 with a movable pointer that *E* could use to display a score to *S* after each trial. The zero end of the scale was to the *S*'s right. A score of 20 was clearly marked "goal score," and the words "Much too high" were placed beyond 40 at the left of the scale. The scale was concealed by a hinged shield except during the feedback period.

Procedure: When a pair of *Ss* was seated in the booths, *E* talked with them briefly to "set them at ease." Then tape-recorded instructions, which were identical for all *Ss*, were played. The instructions described the knob and feedback scale and instructed *S* to turn his knob to the left (counterclockwise) and try to make a score of 20 each time a buzzer sounded. No mention was made of their working together in any way.

Two independent variables were manipulated to conform to a 2 × 2 design by varying the type of feedback displayed to *S* after each trial. During the 25 experimental trials (trials 11–35) all *Ss* received socially confounded feedback; more specifically, they received a team score consisting of a weighted average of their two responses. One independent variable was the relative weighting of the two teammates in this feedback. In condition 1, teammates were equally weighted, i.e.,

$$\text{Feedback} = \tfrac{1}{2} R_1 + \tfrac{1}{2} R_2,$$

TABLE 1

Feedback in each Experimental Condition

Condition	Feedback was: Pretraining Period (Trials 1–10)	Experimental Period (Trials 11–35)
A1	Team score, equal weights	Team score, equal weights
B1	Own response	Team score, equal weights
A2	Team score, 3:1 weights	Team score, 3:1 weights
B2	Own response	Team score, 3:1 weights

where R_1 and R_2 were the response magnitudes of the two *Ss*. In condition 2, one teammate was weighted three times as heavily as the other; i.e.,

$$\text{Feedback} = \tfrac{3}{4} R_1 + \tfrac{1}{4} R_2.$$

The second independent variable was the type of feedback received during pretraining (trials 1–10). In condition A, the *Ss* were pretrained under exactly the same socially confounded feedback that they were to receive during the experimental trials. In condition B, each *S* was pretrained with accurate feedback of his own response magnitude. Table 1 summarizes the manipulations of feedback in each experimental condition. Half the *Ss* in each experimental condition were in AFQT Category II; one quarter were in Category I and one quarter in Category III.

Each trial was timed as follows: (a) a response period of 8 seconds, indicated to *Ss* by a mild buzzer; (b) an 8-second delay before the feedback value was displayed to *S*, during which *E* recorded the response magnitude and made the appropriate setting of the scale pointer (still concealed from *S*); (c) an intertrial interval that varied between 4 and 10 seconds, during which *E* reset the micrometers to zero. The hinged shield covering the feedback scale was raised during the third interval, making the feedback visible to *S only* during this interval. There was no break following the pretraining trials and no change in procedure other than a change in the way *E* computed the feedback value.

RESULTS

Two team measures based upon individual responses were computed for each trial: (a) team accuracy—i.e., the deviation of the team score that had been used as feedback from the "goal score" of 20, expressed as an absolute value; and (b) role differentiation—i.e., the difference between teammates in response magnitude, expressed as an absolute value. Since team accuracy was badly skewed, it was subjected to a logarithmic transformation before statistical analysis. Two separate analyses of variance, including linear trend, were computed on each response measure. Separate

TABLE 2

Trend Analyses of Pretraining Period (Trials 1–10)

Source	Team Accuracy			Differentiation		Error Term
	df	MS	F	MS	F	
1. Teams	31					
a. Pretraining conditions	1	34	0.0, 0.0	2,714	0.8, 4.4*	1b, 3d
b. Teams within conditions	30	4,738		3,201		
2. Trials	9					
a. Linear slope	1	131,095	46.9‡,1 01.5‡	31,079	10.8†, 50.0‡	3c, 3d
b. Residual	8	5,070		2,162		
3. Teams × Trials	279					
a. Between slopes, pretraining conditions	1	372	0.1, 0.3	3,389	1.2, 5.4*	3c, 3d
b. Condition residuals	8	1,500		291		
c. Between team slopes	30	2,797		2,868		
d. Team residuals	240	1,291		622		
Total	319					

*P < .05.
†P < .01†.
‡P < .001.

trend analyses were made of the pretraining period (trials 1–10) and the experimental period (trials 11–35) because of the changes of experimental conditions after trial 10. This type of trend analysis has been described by Alexander (1) and Grant (3).

Table 2 reports the trend analyses of both measures during the pretraining period. In these analyses, differences between weighting conditions are not considered because half of the *Ss* were receiving direct individual feedback. Table 3 reports trend analyses of both measures during the experimental period. In these analyses effects of pretraining, weighting, and their interaction are all analyzed.

Figure 1 shows the performance curves in terms of team accuracy for the two different conditions of pretraining. Since there were only minor differences between weighting conditions, they have been combined in the graph for simplicity. Figure 2 is a graph of the role-differentiation measure. Again the two weighting conditions have been combined for graphical simplicity.

In the trend analyses in Tables 2 and 3, each F ratio is computed with two different error terms. The first test in each case is based upon the error term suggested by Grant (3), and the second is based upon the error term suggested by Alexander (1). The first test is the appropriate one for general-

izing the results across teams and is the more conservative test in each case reported here. The more liberal second test is also included because the research area is relatively novel and such tests may suggest hypotheses for further testing in later research.

During the pretraining period (see Table 2), team accuracy improved markedly, and this increase in accuracy was highly significant (source 2a). However, there were no differences between pretraining conditions in accuracy (source 1a) or rate of improvement (source 3a). During the same period the role-differentiation measure decreased significantly (source 2a). In teams trained on individual feedback, role differentiation was less (source 1a) and decreased at a greater rate (source 3a), though these differences were significant only with the more lenient error term.

During the experimental period (see Table 3), team accuracy continued to improve, though very slowly, but this continued improvement (source 2a) was significant only with the more lenient F test. More important, however, was the fact that teams pretrained on *team* feedback improved at

TABLE 3

Trend Analyses of Experimental Period (Trials 11–35)

Source	Team Accuracy			Differentiation		Error Term
	df	MS	F	MS	F	
1. Teams	31					
a. Weighting conditions	1	2312	0.3, 3.7	281	0.1, 5.1*	1d, 3f
b. Pretraining conditions	1	981	0.1, 1.6	5692	1.8, 103.5‡	1d, 3f
c. Interaction, 1a × 1b	1	715	0.1, 1.2	2	0.0, 0.0	1d, 3f
d. Teams within conditions	28	7755		3079		
2. Trials	24					
a. Linear slope	1	2871	4.1, 4.6*	625	3.9, 11.4†	3e, 3f
b. Residual	23	477		50		
3. Teams × trials	744					
a. Between slopes, weighting conditions	1	1239	1.8, 2.0	46	0.3, 0.8	3e, 3f
b. Between slopes, pretraining conditions	1	3252	4.7,* 5.3*	335	2.1, 6.1*	3e, 3f
c. Between slopes, interaction	1	30	0.0, 0.0	127	0.8, 2.3	3e, 3f
d. Condition residuals	69	899		58		
e. Between team slopes	28	695		160		
f. Team residuals	644	618		55		
Totals	799					

*$P < .05$.
†$P < .01$.
‡$P < .001$.

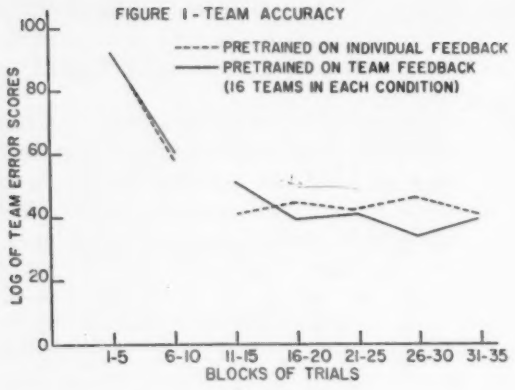


FIG. 1

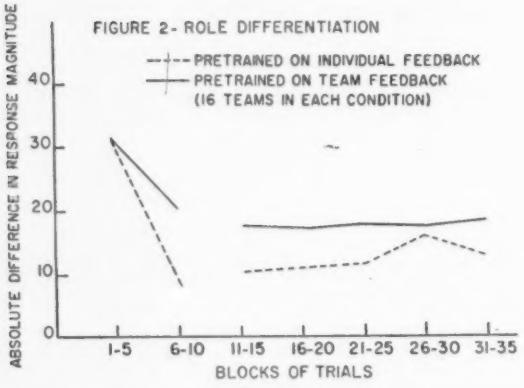


FIG. 2

a faster rate than those pretrained *individually*, and this difference (source 3b) was significant with either error term. On the other hand, the role-differentiation measure during the experimental period reversed its trend from the pretraining period. Role differentiation increased (source 2a), and the rate of increase was greater for those teams pretrained individually than for those pretrained on team feedback (source 3b), though these differences were significant only with the more lenient error term.

Over the 25 experimental trials, role differentiation was greater in teams pretrained on team feedback than those pretrained individually, and slightly greater in teams with unequal weighting than in teams with equal weighting. Neither of these differences (sources 1a and 1b) generalizes over teams, as indicated by the conservative error term, but both are significant with the more lenient error term.

DISCUSSION

It seems apparent that predictions about group performance under socially confounded feedback will have to consider certain properties of the "feedback function" (i.e., the function by which individual responses are combined into a single feedback) and the changes in this function over time. Much empirical and theoretical analysis may be required before all the relevant properties of this function can be identified.

One property of the feedback function that might be expected to be of importance is the relative weight given to the contributions of different teammates. In this study two different weightings were used. Under the conditions of this study, it made little difference whether teammates contributed equally or in a 3 to 1 ratio to the team score that was used as feedback. The only difference noted was that when the teammates' contributions were weighted equally in the feedback, they responded more nearly equally—a difference that was consistent over trials but not over teams.

With the feedback functions used in this study it was possible for one teammate to compensate for another's errors. That is, for example, if one team member was responding too little, the other could increase his response, and both would obtain a satisfactory team score. Another way of stating this condition is to say that the optimal response for any team member depended upon what response his partner made. It would be possible to impose other feedback functions such that there is a fixed optimal response for one team member (or both) regardless of the teammate's response. Indeed, one important property of feedback functions appears to be the possibility for one teammate to compensate for another's errors. This possibility is determined by the feedback function, the range of responses available to each individual, and the magnitude of the goal score. It appears that a feedback function that allows mutual compensation by teammates simply increases the number of combinations of responses that will achieve the desired score, and hence be "reinforced." Thus, the probability of achieving accuracy quickly may be enhanced. As will be discussed below, this gain may be at the cost of having widely different combinations of individual responses in different teams, which entails less flexibility for changing the composition of teams.

Another independent variable in this experiment was a change in the feedback function after the first 10 trials. That is to say, *Ss* were given different experimental histories with respect to this particular task and then were run through 25 experimental trials.

During the initial 10-trial pretraining period, teams working under socially confounded feedback and teams working under individual feedback both achieved about the same level of accuracy at about the same rate. This result seems surprising at first glance since it might seem that indi-

vidual feedback should yield better results. However, this result can be explained by the fact that mutual compensation for errors was possible under the confounded feedback. Since the confounded feedback was a simple weighted average, it allowed teams to achieve the desired team output with a variety of combinations of individual responses. Hence, if one teammate learned slowly and continued to respond too little or too much, the other could compensate for him and obtain an accurate team output. On the other hand, under individual feedback, *both* teammates must approximate their goal scores before the team output is accurate. One might even expect from this reasoning that teams on confounded feedback would achieve any given level of team accuracy *more* quickly than teams on individual feedback.

During the experimental period (trials 11–35), when all teams received confounded feedback, those teams pretrained on confounded feedback improved their accuracy at a faster rate than those pretrained with individual feedback. It is not possible to say with certainty whether the *type* of pretraining accounts for this difference, or whether simply a *change* in the type of feedback retards subsequent improvement. Distinguishing clearly between these explanations would require a differently designed experiment with some additional experimental conditions.

Besides team accuracy, another dependent variable in this experiment was a simple role-differentiation measure—the difference between the response magnitudes of the two teammates. The same team output (and hence team accuracy) could be achieved by various combinations of responses by the separate teammates. The differentiation measure indicates how different the two contributions to the team output were; i.e., how differently teammates responded.

There were marked differences between pretraining conditions in the amount of role differentiation: in teams trained individually the two teammates responded more similarly than in teams trained on team feedback, at least immediately after training. However, when the individually trained subjects started receiving team feedback, their role differentiation increased until it was of about the same magnitude as in teams trained on team feedback from the start.

Role differentiation as a dependent variable may be important to consider in cases where team composition is not constant; i.e., where individuals do not work with the same teammate all the time. Minimal role differentiation would minimize problems of retraining or adapting to new teammates. Pretraining on individual feedback has the advantage of minimizing this role differentiation, at least temporarily, but the advantage appears to be dissipated rapidly when subjects so trained start working under socially confounded feedback.

In conclusion, it may be noted that caution is required in extrapolating the trends of the performance curves in Figures 1 and 2. It is interesting to speculate how long these trends might continue if the number of trials were extended. A likely guess is that eventually the values of both variables would stabilize with no difference between conditions. In other words, the effect of pretraining would probably be dissipated after enough additional trials.

SUMMARY

An experimental study was made of the performance of two-man teams at a simple task under "socially confounded feedback"; i.e., the knowledge of results available to each team member was a single team score that combined response measures of both individuals. Sixty-four experimental subjects (*Ss*) were used, in pairs, and were run through 35 timed trials, receiving a weighted average of their two responses as feedback. The two independent variables, arranged in a 2 × 2 design, were: (a) pretraining (10 trials) under individual vs. confounded feedback, and (b) relative weighting of responses of the two *Ss* in the feedback (1:1 vs. 3:1). The dependent variables were (a) team accuracy and (b) role differentiation (i.e., difference between teammates in response magnitude). Teams pretrained on confounded feedback improved their accuracy at a greater rate than did teams pretrained on individual feedback. Individual pretraining was associated with less role differentiation immediately after training, but this difference diminished during later practice until at the end there were no differences between experimental conditions.

Manuscript received: August 5, 1957

Robert L. Hall
Systems Operator Research Unit
Operator Laboratory Field Unit Number 1
AF Personnel and Training Research Center
P. O. Box 1557
Lackland AFB, Texas

REFERENCES

1. Alexander, H. W., "A General Test for Trend," *Psychological Bulletin*, 1946, 43, 533–537.
2. Gragg, D. B., and H. J. Douglass, *Conversion tables for selected airmen classification battery scores and comparable scores on other selected service and civilian tests*. Lackland Air Force Base, Texas: Air Force Personnel and Training Research Center, June, 1956. (*Development Report* AFPTRC-TN-56-89.)
3. Grant, D. A., "Analysis-of-variance Tests in the Analysis and Comparison of Curves." *Psychologicel Bulletin*, 1956, 53, 141–154.

Intercity Telephone and Airline Traffic Related to Distance and the "Propensity to Interact"[1]

CARL HAMMER, *The Sperry Rand Corporation, New York*

FRED CHARLES IKLÉ, *The RAND Corporation, Santa Monica, California*

The relationship between spatial distance and the frequency of interactions (such as telephone calls, visits, or mail) has been the subject of several investigations in recent years. Best known are the studies by George Zipf (13, 14) John Stewart (9, 10), and Stuart Dodd (4, 5). Related to the same problem are various articles on migration and the many papers on residential propinquity as a factor in mate selection (3). Furthermore, traffic engineers have made similar studies which are usually not so well known among sociologists, and which often suggest the same formulas independently.[2]

These investigations evoked so much interest probably because they led to generalizations which seemed clearly defined, verifiable, and applicable to a great variety of phenomena—conditions not easily met in social science. The pioneers in this area, therefore, were tempted to build "universal" theories on their rather limited statistical evidence. Zipf, for example, concluded—on the basis of an imperfect statistical analysis—that there was a *linear* relationship between the frequency of interactions and distance. Then he proceeded to demonstrate theoretically why this relation had to be linear. However, upon closer scrutiny, neither his analysis nor other studies on this topic prove the existence of a linear relationship.[3] As will be shown below, our evidence suggests rather a *nonlinear* function.

Actually, the rectification of the distance function may not be of great

[1] This paper is based on a research project which was undertaken at the Bureau of Applied Social Research, Columbia University, and assisted by a grant to that University from the Rockefeller Foundation. The authors are indebted to Dr. M. Ray Mickey, The RAND Corporation, Santa Monica, Calif., for valuable suggestions, and to Mrs. Bonnie C. Loftin for her assistance in the analysis of the data. Carl Hammer was affiliated with the Franklin Institute, Philadelphia, while this project was carried out. This article may be identified as publication No. A-240 of the Bureau of Applied Social Research.

[2] See Carrothers (2) for an extensive bibliography of studies dealing with interactions over distance.

[3] Full credit is, of course, due to Zipf for bringing this subject to the attention of sociologists. The reason why many studies erroneously concluded that there was linearity (or some other theoretically postulated relationship) is discussed in Fred C. Iklé's Comment on Anderson's article, *American Sociological Review*, 1955, 20, pp. 713–714.

interest except for those concerned with traffic estimates. The sociologically more interesting findings begin after the influence of distance has been allowed for so that the frequency of interactions can be related to other socioeconomic factors. We will first examine the relationship between frequency of interactions and distance alone, by means of a logarithmic regression. Then we will proceed to ferret out some residual variations that can tentatively be interpreted as differences in the individual "propensity to interact."

The basic principle is simple. Theoretically, we can assign a definite location in space to each person in a given population (at a certain time) and thus define the distance for each possible pairing of persons. Furthermore, we can determine the number of interactions (of a given type) for each pair during a specified time. One expects that, on the average, the number of interactions per pair becomes smaller with increasing distance. The question is which functions will best express this relationship.

However, empirical data on *individual* pairs of people can be collected only on a small scale. But there is a way of obtaining grouped data. We can take two cities and consider all the possible intercity pairings of two people, treating the distance in all these pairs as the same. Thus, we need only two measurements for each pair of cities: (a) the intercity distance and (b) the total number of interactions between the cities, which is the sum of the interactions in all intercity pairs For our purposes interactions should be enumerable and must take place between two persons who are geographically separated. The second condition is not completely met by some of our empirical data; for example, trips do not necessarily represent interactions between two people (6).

Most interactions have a direction: they originate in one place and terminate in the other. Our data on city pairs show that the frequency of interactions in the two directions differs significantly. But for the purpose of the present paper we ignore this and lump interactions in both directions together. It would be possible to extend the analysis to take account of direction. However, the sociological interpretation would lead beyond the scope of this study involving new concepts, such as "originators" vs. "recipients" of interactions.[4] Therefore the formulas developed here make no distinction between initial and terminal points of interactions. The following notation is used:

$i, j = 1, 2, 3, \cdots, n$: Index of the city, where n is the total number of cities studied.

$N = \frac{1}{2}n\,(n-1)$: Number of possible pairings of cities.

P_i = Number of people (interactors) in the ith city.

$D_{ij} \equiv D_{ji}$ = Distance between the ith and jth city (in some convenient units).

[4] This distinction has been made for interactions in small face-to-face groups (1).

$H_{ij} = H_{ji}$ = Number of interactions between the ith and jth cities during a given period, regardless of direction; i.e., the interactions originating in city i and terminating in city j are added to those originating in city j and terminating in city i.

In addition, we define $D_{ii} = H_{ii} = 1$.

THE REGRESSION BETWEEN THE FREQUENCY OF INTERACTIONS AND DISTANCE

There are P_iP_j possible pairings of two persons between city i and city j whose populations are P_i and P_j. Hence, the mean number of interactions per pair of people between city i and city j is $H_{ij}/(P_iP_j)$. This number can be plotted as the dependent variable against distance as the independent one. For certain types of interactions and within a certain range of distance the scatter diagram on log-log paper resembles a straight line. This suggests the function:

$$\log H_{ij} - \log P_i - \log P_j = \log k - b \log D_{ij} \tag{1}$$

or in nonlogarithmic form

$$H_{ij}/(P_iP_j) = kD_{ij}^{-b} \tag{2}$$

where b and k are constants for a given type of interaction and within a given range of distance and time.

Function [1] has been introduced and applied to empirical data by several previous investigators. Usually the constant b has been set equal to 1 so that this function became a simple inverse-linear relation. However, as will be shown below, statistical tests indicate that the exponent b is significantly different from unity. The condition that $\sum (kP_iP_jD_{ij}^{-b} - H_{ij})^2$ be a minimum leads to an implicit transcendental equation, only solvable by iterative methods. Hence, we shall compute this regression with the logarithmic formulation.[5]

Formula [1] leads now to

$$\begin{aligned} \log H_{ij} &= \log k + \log P_i + \log P_j - b \log D_{ij} \\ &= \log k + \log P_i + \log P_j - b\,(\log D_{ij} - \Delta) \end{aligned} \tag{3}$$

where $\Delta = \sum (\log D_{ij})/N$.

Partial differentiation of $\sum [\log k + \log P_i + \log P_j - b(\log D_{ij} - \Delta) - \log H_{ij}]^2$ with respect to $\log k$ and b yields

$$\widehat{\log k} = \sum \log H_{ij}/N - 2 \sum \log P_i/n, \tag{4}$$

$$\begin{aligned} \hat{b} = [\sum_i (\log P_i \sum_j \log D_{ij}) + \Delta \sum \log H_{ij} - (n-1)\Delta \sum P_i \\ - \sum \log H_{ij} \log D_{ij}]/[\sum (\log D_{ij})^2 - N\Delta^2]. \end{aligned} \tag{5}$$

[5] This minimizes the *percentage* error, but there are reasons why one might prefer to minimize the absolute error if this were not so cumbersome.

The partition theorem for the chi-square distribution shows that the estimates for log k and b are stochastically independent. An estimate of the variance is given by

$$[6] \qquad S_1^2 = \sum (\log H_{ij} - \log \hat{H}_{ij})^2/(N-2).$$

The parameters have the estimated variances

$$[7] \qquad \text{Var } (\log k)' = S_1^2/N.$$

$$[8] \qquad \text{Var } (b) = S_1^2/[\sum (\log D_{ij})^2 - N\Delta^2].$$

The estimate of the variance of any computed value log H_{ij} for some log D_{ij} is finally

$$[9] \quad \text{Var } (\log H_{ij}) = S_1^2\{1/N + [\log D_{ij} - \Delta]^2/[\sum (\log D_{ij})^2 - N\Delta^2]\}.$$

The null-hypothesis tested with this model assumes that there is no relation between distance and the frequency of interactions per pair of people. Expressed in logarithms for the parameter log k_0 it is

$$[10] \qquad \log H_{ij} = \log P_i + \log P_j + \log k_0.$$

Application of least squares yields

$$[11] \qquad N \log \hat{k}_0 = \sum \log H_{ij} - (N-1) \sum \log P_i.$$

The variance S_0^2 is obtained from

$$[12] \qquad (N-1)S_0^2 = \sum [\log H_{ij} - \log P_i - \log P_j - \log \hat{k}_0]^2.$$

Introducing City Weights into the Regression between Interaction and Distance

The social and economic compositions of city populations differ substantially. This affects the number of interactions between city pairs, because the frequency of interactions over a given distance between two persons depends, among other factors, on their social and economic characteristics. One city might contain predominantly people whose propensity to interact is low, and conversely, the inhabitants of another city might tend to interact frequently. One can express and empirically measure these differences in various ways. One convenient form is to introduce multiplicative weights for the city populations into our equation [2]:

$$[13] \qquad H_{ij} = k(P_iW_i)(P_jW_j)D_{ij}^{-b}$$

where W_i, W_j are the city weights. Solution of the maximum likelihood equations for this model again yields unwieldy, implicit transcendental

equations which can only be solved by iterative methods. Therefore, the corresponding logarithmic form will be employed:

$$[14] \quad \log H_{ij} = \log k + \log P_i + \log W_i + \log P_j + \log W_j - b(\log D_{ij} - \Delta).$$

The solution of the corresponding maximum likelihood equations shows that not all parameters are stochastically independent. We find

$$[15] \quad \log \hat{k} = \sum \log H_{ij}/N - 2 \sum \log P_i/n \sum_i [(\sum_j \log H_{ij})(\sum_j \log D_{ij})]$$

$$[16] \quad \hat{b} = \frac{-n\Delta \sum \log H_{ij} - (n-2) \sum \log H_{ij} \log D_{ij}}{(n-2) \sum (\log D_{ij})^2 + nN\Delta^2 - \sum_i (\sum_j \log D_{ij})^2}$$

$$[17] \quad \log W'_\lambda = \{\sum_j \log H_{\lambda j} - (n-1) \log k - (n-2) \log P_\lambda - \sum \log P_i + b [\sum_j \log D_{\lambda j} - (n-1)\Delta]\}/(n-2).$$

The partition theorem for the chi-square distribution shows that an estimate of the variance is obtained by

$$[18] \quad S_2^2 = \sum (\log H_{ij} - \log \hat{H}_{ij})^2/(N - n - 1).$$

The estimated variance of $\log k$ is given by

$$[19] \quad \text{Var } (\log k) = S_2^2/N.$$

The estimates $[\hat{b}, \log \hat{W}_i]$ are normally correlated with means $[b, \log W_i]$. The estimate of the variance of b is

$$[20] \quad \text{Var } (\hat{b}) = S_2^2 \Big/ \left\{ \sum_{i<j} (\log D_{ij} - \Delta)^2 - \frac{1}{n-2} \left[\sum \delta_i^2 - \frac{(\sum \delta_i)^2}{n} \right] \right\}$$

where $\delta_{ij} \equiv \sum_{j \neq i}^{j=n} \log D_{ij}$.

Empirical Data on Telephone Calls

Data on the number of telephone calls between two cities have the advantage that they clearly represent interactions between *two persons*, which is a prerequisite for our model. We used figures on the number of telephone calls during the month of October in the years 1946, 1948, and 1952, between 10, 24, and 23 cities respectively. In accordance with our model, calls in both directions were combined. For a few possible pairings of cities no data were available and our formulas were adjusted correspond-

TABLE 1

Parameters Obtained from the Various Hypotheses

Hypothesis or Model	Parameter	Telephone Calls			Airplane Trips March 1950
		Oct. 1946	Oct. 1948	Oct. 1952	
Null	S_0^2	0.207	0.247	0.247	0.421
Logarithmic without city weights	S_1^2	0.033	0.089	0.078	0.264
	log k	−7.061	−7.425	−8.041	−9.603
	b	1.001	1.136	1.283	1.204
Logarithmic with city weights	S_2^2	0.009	0.025	0.023	0.016
	log k	−7.061	−7.424	−8.041	−9.603
	b	1.471	1.638	1.690	1.743
Number of cities		10	24	23	27
Number of city pairs with observed data*		43	266	238	351

Note: The 95 per cent confidence intervals for log k lie within less than 1 per cent and those for b within less than 15 per cent above and below the best estimate.

* Data not available for all city pairs.

ingly. As the populations, P_i, of the cities we took the number of telephone subscribers.[6] As distances, D_{ij}, we used the shortest airline distance between the cities (12). Many alternative measures of distance could, of course, be suggested. Incidentally, using the *cost* of telephone calls instead of distance leads to less satisfactory results, primarily because rates increase in a few big steps.

Table 1 lists the estimates for the parameters in the model with city weights and in the one without. Table 2 shows the values obtained for the city weights (in antilogs). A weight larger than unity indicates an average propensity to interact (in this case, to send out or receive telephone calls) in the given city which is above the average of all cities; and conversely for a weight below unity.

It is desirable to arrive at estimates of the city weights independent of the number of observed interactions. If the weights could only be determined a posteriori on the basis of observed calls, it would be impossible to use them for estimating interaction frequencies for cities about which no interaction data are available. However, the weights could still be used to estimate frequencies for occasional city pairs that are without data—a useful application for scheduling an airline between cities previously not connected!

[6] The areas (so-called toll areas) from and to which the calls have been counted are usually somewhat larger than the political city. The number of subscribers applies to these areas, as it should in conformity with our model. Source for the number of subscribers: U. S. Federal Communications Commission, *Statistics of the Communications Industry, passim.*

TABLE 2

City Weights

City	Telephone Calls			Airplane Trips March 1950
	Oct. 1946	Oct. 1948	Oct. 1952	
Akron, O.	—	0.56	—	—
Atlanta, Ga.	—	—	0.91	1.29
Baltimore, Md.	0.70	0.60	0.64	0.14
Boston, Mass.	—	1.41	1.58	0.53
Buffalo, N. Y.	—	0.68	—	—
Chicago, Ill.	1.12	1.07	1.18	1.13
Cleveland, O.	0.76	0.68	0.61	0.73
Dallas, Tex.	—	—	1.22	1.76
Denver, Colo.	—	1.16	—	1.72
Detroit, Mich.	0.69	0.60	0.56	0.59
Flint, Mich.	—	0.34	—	—
Hartford, Conn.—Springfield, Mass.	—	—	—	0.33
Houston, Tex.	—	1.30	—	1.21
Kansas City, Mo.	—	—	—	1.38
Los Angeles, Calif.	3.00	2.94	3.06	1.82
Louisville, Ky.	—	—	—	0.60
Memphis, Tenn.	—	1.14	—	0.93
Miami, Fla.	—	2.58	—	7.61
Milwaukee, Wis.	—	0.52	0.59	—
Minneapolis, Minn.	—	0.74	0.85	0.94
Nashville, Tenn.	—	—	—	0.80
Newark, N. J.	—	2.04	2.03	—
New Orleans, La.	—	1.22	0.94	1.22
New York, N. Y.	1.20	1.03	1.01	0.84
Norfolk, Va.	—	1.42	—	—
Omaha, Neb.	—	—	0.88	—
Philadelphia, Pa.	0.74	0.62	0.65	0.24
Phoenix, Ariz.	—	—	2.03	2.85
Pittsburgh, Pa.	0.70	0.67	—	—
Portland, Ore.	—	—	1.62	1.15
Providence, R. I.	—	—	—	0.20
Richmond, Va.	—	—	0.58	—
San Francisco, Calif.	—	1.87	—	2.34
St. Louis, Mo.	1.02	1.12	1.04	0.66
St. Paul, Minn.	—	—	0.61	—
Salt Lake City, Utah	—	—	0.93	—
Seattle, Wash.	—	—	1.65	2.08
Tampa, Fla.	—	—	—	4.21
Tulsa, Okla.	—	—	0.81	—
Washington, D. C.	1.17	1.09	—	—

Note: The 95 per cent confidence intervals lie about 15 per cent below and above the best estimate.

TABLE 3

Rank Correlations between City Weights and Various Socioeconomic Characteristics of the Cities

Characteristics According to Which Cities Have Been Ranked*	Direction and Degree of Correlation†			
	Telephone Calls			Air Trips
	1946	1948	1952	1950
1. Number of managers, proprietors, and officials as a percentage of total city population	+++	++	++	++
2. Annual retail sales in city divided by city population	+	+	+	+
3. Number of rooms for transients in the Metropolitan Area divided by Metropolitan population	++	+	+	+
4. Percentage of nonwhites in Urbanized Area	++	+	+	−
5. Median income of families and unrelated persons in Urbanized Area	−	−	−	−
6. Percentage of families having incomes of $5000 or more in city	+	−	+	−

* Sources: U. S. Bureau of the Census, *County and City Data Book 1952*, pp. 442–505; 1950 census of population; and 1948 census of business. Data apply to 1950 for row 1 and 4, to 1948 for row 2 and 3, and to 1949 for row 5 and 6.

† Positive, +, (or negative, −) correlation; where ++ (or −−) means significant at 95 per cent level, and +++ (or −−−) means significant at 99 per cent level.

We calculated rank correlations between the city weights and various socioeconomic characteristics of the cities which might presumably account for different propensities to interact (Table 3). In addition, the city weights are correlated with the average distance of a city from all other cities. Far-away cities generally have higher weights (some notable exceptions are Newark and Boston which have high weights but low average distances).

Empirical Data on Airline Trips

Airline trips between cities in the United States offer another example of interactions about which adequate statistics are available to test our hypotheses. However, as already mentioned, a plane trip by a passenger does not necessarily represent an interaction between a person at the place of origin (the passenger) and a person at the place of destination (e.g., the person visited). But we can assume that the number of person-to-person interactions is roughly proportional to the number of trips. (If it were exactly proportional then it would be immaterial whether we used the number of actual interactions or the number of trips.)

A second disadvantage of airline data is that the number of potential interactors, P, is less exact than in the case of telephone calls, because it

is difficult to estimate the area from which an airport draws its passengers. We used the population of the urbanized area (as defined by the 1950 census) that includes the city in which the airport is located. The period during which all passenger trips on scheduled flights have been counted is the month of March 1950 (12). The intercity distance is again the shortest airline distance.

Table 1 lists the values obtained for the parameters of our formulas. The computations have been carried out for a group of 27 cities (351 city pairs). Table 2 presents the values obtained for the city weights. It becomes apparent that the traffic situation of the city plays an important role in these weights. For example, if there is another airport nearby with better schedules it tends to draw traffic away from a city, as probably is the case of Baltimore with its proximity to the National Airport in Washington. This question of special traffic situations has been examined in other studies (11). In some cities airline traffic may be lower because other important cities are located so closely that railroad transportation becomes preferable; for example Philadelphia. The weights for Miami and Tampa, Florida, are probably so high because of vacation traffic (month of March!). Such trips do not represent interactions between two people and hence are extraneous to our model. Table 3 gives the rank correlations between the city weights and various socioeconomic characteristics of the cities.

SUMMARY OF FINDINGS

1. A statistically significant relationship exists between distance and the frequency of telephone calls or airline trips per pair of persons, between cities in the United States. In theory, the statistical tests developed here are applicable to any other type of interaction which takes place between two persons who are geographically separated.

2. The frequency of interactions between two persons can be expressed as a power function of distance, with a negative exponent, $-b$, where b assumes various values between confidence limits of 1.3 and 1.8 in the case of telephone and airline traffic as interactions. These confidence limits indicate that all the data used in this study fail to justify either the inverse linear or the inverse square "law," which previous investigators had suggested for the distance function.

3. The fact that the available empirical data are grouped into city pairs introduces an additional variation due to the particular population composition or some other peculiarities of the cities. This variation can be expressed and measured with multiplicative weights for each city population. The "city weights" were found to be correlated positively and consistently for all sets of data (telephone and airline) with:

a. number of managers, officials, and proprietors as a percentage of the cities' total populations
b. the annual retail sales divided by the city population
c. the number of rooms for transients in the Metropolitan Area divided by total Metropolitan population.

Somewhat perplexing is the negative correlation of the city weights with median family income, which implies that the higher a city's median family income the lower its per capita number of interactions (telephone calls or plane trips). There may be intercorrelations with other factors which account for this fact. Further research will be needed to explore the relationship between city weights and socioeconomic characteristics of the cities.

The city weights can be interpreted as an index of the average "propensity to interact" of all individuals in the city.[7] This reminds us of interaction matrices from small group studies, such as the one published by Bales and associates (1), which can also be factored out so that the interaction propensity of individual A multiplied with the interaction propensity of individual B gives the number of interactions between A and B (7).

4. The mechanism by which spatial distance is related to the frequency of human interactions is very involved and cannot be explained by some simple sociologic or economic interpretation. The cost of interacting over distance is obviously a factor which reduces the frequencies as distance increases. But there are interactions with a constant cost regardless of distance, which also become less frequent the farther apart people live, e.g., mail! Furthermore, distance affects the probability that a relationship between two persons will become established.

5. The fact that one can find a simple, statistically significant relationship between interactions, distance, and city characteristics is definitely useful for communication and traffic studies. It can be applied to estimate

[7] John Q. Stewart and Stuart C. Dodd hypothetically suggested such multiplicative weighting factors. To our knowledge this is the first time that numerical values have been determined for them (except in small group studies). It must be admitted, however, that our city weights partly reflect the geographical location or other special situations of the city (such as poor train or airline connections) which have nothing to do with the "interaction propensity" of the inhabitants. But in most cities more interesting differences can be found that seem to be due to the people and not merely to the logistics of traffic. Note, for example, that people in Detroit took only about one half as many airplane trips as people in other U.S. cities, and made only about two thirds as many phone calls, whereas the Los Angelites made three time as many phone calls and took almost twice as many plane trips (more out-of-town connections because of recent migration?). Conversely, Bostonians made about one and one-half times as many phone calls but took only half as many plane trips as the average.

intercity traffic for highway construction and to evaluate airline schedules.[8] Furthermore, if one knows how spatial distance affects interactions between people one can then account for this effect of distance and analyze residual deviations.

Manuscript received: March 15, 1957
Revised manuscript received: August 2, 1957

Fred C. Iklé
RAND Corporation
1700 Main Street
Santa Monica, California

REFERENCES

1. Bales, R. F., F. L. Strodtbeck, T. M. Mills, and M. E. Roseborough, "Channels of Communication in Small Groups," *American Sociological Review*, 1951, 16, 461–468.
2. Carrothers, G. A. P., "An Historical Review of the Gravity and Potential Concepts of Human Interaction," *Journal of the American Institute of Planners*, 1956, 22, 94–102.
3. Clarke, A. C., "An Examination of the Operation of Residential Propinquity as a Factor in Mate Selection," *American Sociological Review*, 1952, 17, 17–22.
4. Dodd, S. C., "The Interactance Hypothesis, a Gravity Model Fitting Physical Masses and Human Groups," *American Sociological Review*, 1950, 15, 245–256.
5. Dodd, S. C. "A Measured Wave of Interracial Tension," *Social Forces*, 1951, 24, 281–289.
6. Iklé, F. C., "Sociological Relationship of Traffic to Population and Distance," *Traffic Quarterly*, 1954, 8, 123–136.
7. Keller, J. B., "Comment on 'Channels of Communication in Small Groups,'" *American Sociological Review*, 1951, 16, 842–843.
8. Mayhill, G. R., "A Critique of the C.A.A. Studies on the Air Traffic Generation in the United States," *The Journal of Air Law and Commerce*, 1953, 20, 158–177.
9. Stewart, J. Q., "A Measure of the Influence of a Population at a Distance," *Sociometry*, 1942, 5, 63–71.
10. Stewart, J. Q., "Potential of Population and Its Relationship to Marketing," in R. Cox and W. Alderson (eds.), *Theory in Marketing*, Chicago; Richard Irwin, 1950.
11. U. S. Department of Commerce, Civil Aeronautics Administration, *Airline Passengers* (rev.), Washington: U. S. Government Printing Office, 1951.
12. U. S. Department of Commerce, Civil Aeronautics Board, *Airline Traffic Survey* (March, 1950), Washington: U. S. Government Printing Office, 1950.
13. Zipf, G. K., "The P_1P_2/D Hypothesis on the Intercity Movement of Persons," *American Sociological Review*, 1946, 10, 677–687.
14. Zipf, G. K., *Human Behavior and the Principle of Least Effort*, Cambridge, Mass.: Addison Wesley Press, 1949, chap. 9.

[8] The application to the routing and licensing of airlines and for airport planning is particularly important, as shown by the previous research effort by D'Arcy Harvey and others at the Civil Aeronautics Administration (8, 11).

Cumulative Scaling as a Routine Procedure with the IBM 101[1]

EDGAR F. BORGATTA, *Russell Sage Foundation and New York University*

With the development of machines that can both be substituted for clerks and shorten the time required for a given process, many practical problems that face the researcher in treating his data can be alleviated. The availability of machines, however, may influence the actual ways in which data analysis is done, as well as the philosophy underlying the analysis. This paper deals with the machine analysis of cumulative scales.

Using the model of Guttman's "perfect scale," many approaches have been developed and outlined for fitting imperfect data to the model, including those of Guttman (3, 4, 5), Ford (1), Suchman (9), Stouffer and associates (8, 10), Green (2), Lazarsfeld (6), and Marder (7). Except for the approach of Lazarsfeld, which essentially provides another model to fit the same class of data, the techniques may be viewed as assisting the selection of items or providing a series of procedures that allow the researcher economy in his screening of possible arrangements of items. Each of these procedures has advantages and disadvantages, but a limitation common to all is that they are not designed for routinized use.

The following shall outline briefly procedures for routine use with the IBM 101 Electronic Statistical Machine,[2] which is a counter-sorter with a printing device. The special virtue of the IBM 101 is that it has a print and sort impulse that follows the card reading time, thus making it possible to construct logical paths during card reading time in the 60 recode selectors in the machine, and then sending the sort and print impulses through these paths. With 60 recode selectors available, it is conceivable that one could arrange a selection procedure that involves 60 binary decisions, i.e., 60 decisions of whether or not a condition is satisfied. However, because of the limitation of exits for information, as well as the number of selectors that may be used in parallel for any decision, such a selection procedure would be restricted to a small number of paths. On the other hand, using

[1] This report results from recurrent requests for advice on how to do scaling by machine. No claims are made for the originality of the procedures, but certainly credit is due to Samuel A. Stouffer who was concerned with the effective adaptation of the IBM 101 to scaling problems, and to David G. Hays who did much of the "pioneer" learning of the machine. Use of these and related control panels begins with experience at the Harvard Laboratory of Social Relations in 1951 as far as can be determined.

[2] The manual for the IBM 101 (11) is available through salesmen, who are ordinarily more than willing to provide considerable information about the uses and availability of IBM equipment.

a full binary expansion, it is possible to provide all possible arrangements from 7 items. In fact, 33 selectors in the IBM 101 would be required to provide this set of 128 paths (2^7). Again, it would not be possible to exit the machine with all of these since there are only 60 counters, which may be taken as one limitation on exit. It is possible, however, to illustrate how a larger number of exits may be created quite simply. For example, if one wishes to use 6 items on a full binary system and to exit with 64 arrangements (2^6), it is possible to utilize the capacity of the machine in the following way: Count 59 of the arrangements, count the total of the remaining 5 arrangements, and sort 59 arrangements into one pocket and the remaining 5 arrangements into separate pockets. If the remaining 5 arrangements are chosen to be very unlikely events, the burden of hand-counting the cards is minimal since there will either be few or none in each of the 5 sort pockets. Thus, it is seen that the IBM 101 may be utilized to count and sort an imperfect Guttman-type scale of 6 dichotomous items and 64 arrangements. The IBM 101 will do this in one pass of the cards, at the rate of 450 cards per minute, the print cycle taking approximately 17 seconds. To this time should be added the control panel wiring that is involved in creating the 6 dichotomies in terms of the card code on a standardly wired board that would presumably be held as a permanent service board in a machine installation. The time required to wire these 6 dichotomous conditions should be a matter of 2 or 3 minutes. Wiring the entire board, the bulk of which then becomes permanent, should be a matter of 1 or 2 hours if few or no mistakes are made in wiring. To illustrate the principle involved in this kind of wiring, a sample control panel has been drawn for 5 items, with 32 arrangements to be counted. In addition, a total count is taken. On this sample control panel no sort is made. Extension from 5 items to 6 items is a relatively simple matter, but is not feasible in the diagram because of the number of crossing lines to be read. The job that is done by this control panel is indicated in Table A.[3]

In summary, the first control panel we have outlined will allow a person in one pass of the cards to examine a scale of dichotomous items. The question still remains, however, of which items are to be included in such a scale. The answer must be tied to the kind of items available. If, for example, all items are dichotomous to begin with and all are aligned in terms of "positive" meaning in the dimension they are presumed to occupy, then all that is required is that the marginal frequencies for the items be

[3] The diagrams discussed in this paper are available from the American Documentation Institute. Order Document No. 5321 from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C., remitting $1.25 for 35 mm microfilm or $1.25 for 6 by 8 photocopies. Make check or money order payable to: Chief, Photoduplication Service, Library of Congress.

TABLE A

1.	+	+	+	+	+	9.	+	−	+	+	+	17.	−	+	+	+	+	25.	−	−	+	+	+
2.	+	+	+	+	−	10.	+	−	+	+	−	18.	−	+	+	+	−	26.	−	−	+	+	−
3.	+	+	+	−	+	11.	+	−	+	−	+	19.	−	+	+	−	+	27.	−	−	+	−	+
4.	+	+	+	−	−	12.	+	−	+	−	−	20.	−	+	+	−	−	28.	−	−	+	−	−
5.	+	+	−	+	+	13.	+	−	−	+	+	21.	−	+	−	+	+	29.	−	−	−	+	+
6.	+	+	−	+	−	14.	+	−	−	+	−	22.	−	+	−	+	−	30.	−	−	−	+	−
7.	+	+	−	−	+	15.	+	−	−	−	+	23.	−	+	−	−	+	31.	−	−	−	−	+
8.	+	+	−	−	−	16.	+	−	−	−	−	24.	−	+	−	−	−	32.	−	−	−	−	−

34. Total count. Sort into pocket 7, sort selection 2.

Items illustrated in diagram are, ordered from 1 to 5, defining the positive side of the dichotomy: Column 6, punch 0; Column 10, punch 0; Column 5, punch 0 or 1; Column 8, punch 0, 1, or 2; and Column 9, punch 4. A control panel may be wired to handle from 2 to 6 items, or to examine two 4 item scales or four 3 item scales simultaneously.

found. Since only one pass of the cards is required to examine the scale, it is apparent that an efficient procedure is trial and error, based on selection of items of approximate equal distance between the marginals. In terms of our example, it may be strategic to run 5 items that can be selected on the basis of some intuitive or logical expectation that they will be the best ones at the equally spaced positions in terms of the marginal frequencies. If the scale that results on this trial is inadequate or appears to have a high proportion of errors associated with a given item (or with given items), on the next trial substitute another item for the poor one, assuming that another item exists at approximately that marginal frequency. Since the ordering of scales is tied to the ordering of marginal frequencies, the number of trials in this trial-and-error procedure are actually limited. The researcher should be oriented toward finding the best set of items, and then finding what the best number of items in the scale is. It may be that he will find a superior 3-item scale, a good 5-item scale, and a fairly poor 6-item scale. He will have to make his own decision of which scale to retain for his purposes.

The suggestion that a trial-and-error procedure be utilized should not be interpreted as a substitution for the skills required in scaling and related numerical procedures. Rather, this is a suggestion that if it should happen that the items are good enough in regard to the scaling problem, the procedure is economical. Simply stated, if a scale is found in this way there is no reason for proceeding to more elegant procedures.

Researchers are not often satisfied merely by being able to find a scale; they often want to know more about the items that are involved in them. Further, as a matter of routine, researchers may wish to do their scaling on the basis of what we might call "sophisticated" trial and error rather than on the arbitrary trial and error indicated immediately above. The logical way to proceed in this case appears to be to examine the relationship of each item with each other item. For this purpose the IBM 101 is also admirably

TABLE B

Pass one. Split wire to first item. Sort on first item, count total in each position.
Pass two. Split wire to second item. Sort on second item, count on each of five subdecks on item 1 the total in each position.
Pass three. Split wire to third item. Sort on third item, count on each of five subdecks on item 2 the total in each position.
And so on, progressively.
After each sort, intersperse 9 Master cards to cause printout at a, b, c, d, and e, which correspond to the subdecks created by the sorts.

Pass	Total	Sort Count Item	Count											
			Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10	Item 11	Item 12
1	x	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxx	xxx	xxx
2a	x	xxxxx	x	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxx	xxx	xxx
b	x	xxxxx	x	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxx	xxx	xxx
c	x	xxxxx	x	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxx	xxx	xxx
d	x	xxxxx	x	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxx	xxx	xxx
e	x	xxxxx	x	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxx	xxx	xxx
3a	x	xxxxx	xxxxx	x	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxx	xxx	xxx
b	x	xxxxx	xxxxx	x	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxx	xxx	xxx
c	x	xxxxx	xxxxx	x	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxx	xxx	xxx
d	x	xxxxx	xxxxx	x	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxx	xxx	xxx
e	x	xxxxx	xxxxx	x	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxx	xxx	xxx
And so on to last pass. Sort on last pass not necessary; leave in last position (item 12).														
13a	x	x	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxx	xxx	x
b	x	x	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxx	xxx	x
c	x	x	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxxxx	xxx	xxx	x

Items illustrated in the diagram are columns 5 to 13, punches 0 to 4, and columns 14 to 16, punches 0 to 2.

adapted. Table B (corresponding to a second sample diagram) illustrates a cross-tabulation of 12 items, 9 of which have 5 response possibilities, and 3 of which have 3 response possibilities. In addition, as a matter of convenience, space has been allotted to sorting and counting one of the items separately. As indicated, in the first pass of the cards, the total deck is counted and the deck is simultaneously sorted on the first item. It will be noted that this control panel allows a person to count on 60 positions, and thus it can fit a wide variety of situations. For example, it can obviously be used to examine 29 dichotomous items in one series of passes. The fourfold tables resulting from such a series of passes could then be utilized to compute fourfold point correlations or tetrachoric correlations. If the variables have been precoded according to the cutting points necessary for use with Flannigan's coefficient, such a panel may be utilized for examination of a square matrix of 57 items. If time and the number of passes on the IBM 101 are for some reason restricted, it is possible to examine up to 58 dichotomous variables through the knowledge of the totals and by printing of only one side of the distribution in the fourfold table (to be subtracted from the totals for the other side of the fourfold table). Generally, the cross-tabulations present grouped data for correlation procedures as well as for examination at an intuitive level. Such examination of a series of cross-tabulations may be the basis for selection of an initial nucleus of items to be tested in a scale as indicated above or, actually, a routine screening procedure may be utilized.

While the matrix of cross-tabulations may provide the raw material for item analysis, it is possible to indicate another procedure from the item analysis area. This procedure has been noted previously (8) and here we provide an additional routine method for application. In particular, while there may be many relationships involved in a matrix such as created by the procedure outlined in Table B, more than one kind of content may be

TABLE C

Suppose items coded in columns 5, 6, 8, 9, 10 and 12 appear to be highly related, with item 9 obviously stated negatively, and thus the upper half of a dichotomy is indicated (suppose) as follows:

Column	Punches to be Scored "Positive"	
5	0, 1, 2	In one pass through the IBM 101, the cards are sorted according to the number of "positive" responses, from 0 to 6. A count for each sorted value may be wired in paralliel.
6	0	
8	0, 1, 2	
9	4, 3	
10	0, 1	
12	0, 1, 2, 3	

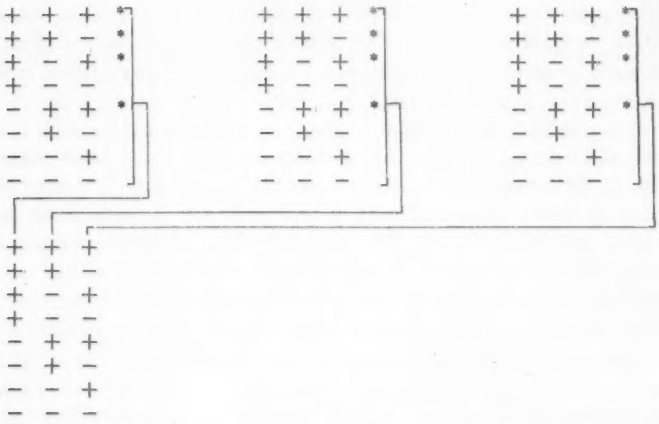
involved. Assuming that a cluster of items may be selected that appears to be highly interrelated, it is then possible to build a short or provisional scale against which all other items are examined. Such a scale could be a rough one, created by a series of items that are dichotomized close to their medians and given simple unit weights to produce an additive score. Table C (corresponding to a third sample diagram) indicates the procedure that is involved, which is of adding the number of "positive" responses that are made, assuming this is the arbitrary way of scoring. Table C indicates the way a score might be produced from a set of 6 items. The practical limits of such a control panel are the available 12 sort pockets and the reject pocket, and, generally speaking, a 13 point range is probably more than adequate for this arbitrary scoring procedure. Once the cards have been sorted according to score, it is possible to run this distribution against the individual items in a control panel such as is illustrated in Table B above. This would give the cross-tabulation of the score with the items and would allow the investigator to judge the relationship of each item in terms of the common content represented in the preliminary score.

On the other hand, Guttman's scaling should be considered as only one alternative in the creation of measures. For all practical purposes this method of item analysis is equally relevant to Likert-type scaling and to scaling according to any other arbitrary cumulative principle. We shall not elaborate the point, but scoring imperfect scale types by the least error criterion and by the middle placement criterion is a close cousin to adding the number of positive responses in the scale pattern of an individual. A researcher will want seriously to explore the relative advantage of Guttman scaling vs. simple additive procedures.

There are several ways of improving scales if a large number of items are available. One way is simply by extending the length of the test, which is a common procedure with Likert-type scales. Another approach that has been suggested has been that of H-technique scaling (8). Basically the logic of the H-technique is that if single items are not discriminating a point on the continuum well enough, then if additional items are available the cutting point may be determined utilizing a number of items and therefore maximizing the common content in regard to the scale at the cutting point. Items are ordinarily grouped according to similar marginals in creation of the H-technique type derived items. Table D presents the scale example which is diagrammed. The example diagrammed does not include a detail matching of emitted punches to column punches since the principle for this is already illustrated in the Table A diagram. The Table D diagram, however, does include an indication of how timing may be utilized in order to set up parallel sets of paths that may be used for both sort and count impulses. Basically the principle involved is that if punches used in the card

TABLE D

Original Items



H-technique scale. Items are scored positive if two or three positives exist in the corresponding three original items, as indicated by asterisk. The illustration in the diagram is to both sort and count the H-technique scale types.

code are in the range from 0 to 9, then the X and Y punches may be utilized for control purposes.[4] In the example, the paths to create the H-technique items are set up by the impulses in the range from 0 to 4; then an impulse is sent through at the time that the X impulse would normally be read on the card and this is utilized to set up a new set of paths corresponding to the 3 H-technique items. Through the new set of paths the sort and count impulses are separately sent, the sort and count impulses coming after card reading time. Although many possible arrangements may be made in the utilization of the H-technique, experience, along with some informal logic, suggests the utility of restricting the operation to H-technique items that are composed of 3 original items each. Since no shortage of selectors exists, the length of the H-techinque scale that can be set up in the IBM 101 control panel is essentially the same as that of Table A. In other words, with minor limitation it is possible to examine in one pass of the cards an H-technique scale with 6 items, based on 18 original items. It is

[4] Persons who utilize IBM equipment are cautioned to plan their data codes to be single-punched ordinary numerical codes. If for some reason the user wants to double-punch a column, let him do it by creating two decks and merging them in a reproducer, so that single-column versions of the cards will be available. IBM computing and tabulating equipment is generally made to handle only one punch in a column, the X and Y punches being controls, and the researcher should leave himself the option for this kind of data treatment.

more likely, however, that on a standard board in an installation, the more appropriate model is an H-technique scale of 4 or 5 derived items.

In summary, we have presented in tabular form 4 kinds of operations that can be done on the IBM 101. Diagrams corresponding to these tables have been prepared to illustrate simple examples. The general utility of the control panels is briefly indicated.

The availability of electronic equipment provides researchers with a considerable saving of time by making complex procedures routine. In addition, the researcher may have greater autonomy in being able to exercise a more direct and greater control over the treatment process. Extensive and intensive data treatment should be expected in terms of any research that is done, and the ability to do scale construction should be recognized as both a routine and minor skill that all investigators in the relevant social and psychological areas should possess. Essentially, the development of equipment to assist in the data treatment processes places a greater responsibility upon the researcher to have the technical skills of his profession since these are tied directly to the possibility of systematically and fully exploiting the research data.

Manuscript received: August 12, 1957

Edgar F. Borgatta
Russell Sage Foundation
505 Park Avenue
New York 22, N. Y.

REFERENCES

1. Ford, R. N., "A Rapid Scoring Procedure for Scaling Attitude Questions," *Public Opinion Quarterly*, 1950, 14, 507–532.
2. Green, B. F., "A Method of Scalogram Analysis Using Summary Statistics," *Psychometrika*, 1956, 21, 79–88.
3. Guttman, L., "The Cornell Technique for Scale and Intensity Analysis," *Educational and Psychological Measurement*, 1947, 7, 247–279.
4. Guttman, L., "The Israel Alpha Technique for Scale Analysis," in M. W. Riley, J. W. Riley, Jr., and J. Toby, *Scale Analysis*, New Brunswick, N. J.: Rutgers University Press, 1954.
5. Guttman, L., "The Quantification of a Class of Attributes: A Theory and Method of Scale Construction," in P. Horst *et al.*, *The Prediction of Personal Adjustment*, New York: Social Science Research Council, Bulletin Number 48, 1941.
6. Lazarsfeld, P. F., "The Logic and Mathematical Foundation of Latent Structure Analysis," in S. A. Stouffer *et al.*, *Measurement and Prediction*, Princeton: Princeton University Press, 1950.
7. Marder, E., "Linear Segments: A Technique for Scalogram Analysis," *Public Opinion Quarterly*, 1952, 16, 417–431.
8. Stouffer, S. A., E. F. Borgatta, D. G. Hays, and A. F. Henry, "A Technique for Improving Cumulative Scales," *Public Opinion Quarterly*, 1952, 16, 273–291.

9. Suchman, E. A., "The Scalogram Board Technique," in S. A. Stouffer *et al.*, *Measurement and Prediction*, Princeton: Princeton University Press, 1950.
10. Toby, J., and M. L. Toby, "A Method of Selecting Dichotomous Items by Cross-Tabulation," in M. W. Riley, J. W. Riley, Jr., and J. Toby, *Scale Analysis*, New Brunswick, N. J.: Rutgers University Press, 1954.
11. IBM, *Electronic Statistical Machine Type 101: Principles of Operation* (Third revision), International Business Machines Corporation, 590 Madison Avenue, New York 22, New York, 1953.